```
=> d his

     (FILE 'HCAPLUS' ENTERED AT 09:08:19 ON 07 APR 1998)
                DEL HIS Y
L1           100 S MAGE 1
L2             0 S 11540
L3             5 S 11540/AB
L4             0 S L3 AND (ANTIBOD? OR MONOCLONAL OR ANTIBOD?/AB OR MONOCL
L5        122515 S PROTEIN# (L) SEQUENCE#
L6        204326 S SEQUENCE#
L7            13 S L1 (L) L6
L8         16383 S TUMOR (L) (PROTEIN# OR ANTIGEN#)
L9          8675 S MELANOMA#
L10        76803 S MOLECULAR WEIGHT OR KD OR KILODALTON? OR KILO DALTON# O
L11       368491 S (MOLECULAR WEIGHT OR KD OR KILODALTON? OR KILO DALTON#
L12          666 S L8 AND L9
L13           64 S L12 AND L11
L14           77 S L12 AND (L10 OR L11)
L15          735 S (46 OR 34) (L) (L10 OR L11)
L16            1 S L14 AND L15
L17           17 S (46 OR 46/AB OR 34/AB OR 34) (L) (KDS OR KDS/AB)
L18            0 S L17 AND L14

=> d .ca l7 1-13; all l16


L7   ANSWER 1 OF 13  HCAPLUS  COPYRIGHT 1998 ACS
AN   1997:625618  HCAPLUS
DN   127:304120
TI   Recombinant tumor-specific antigen-encoding adenoviral vectors for
     human tumor gene therapy
IN   Boon, Falleur; Duffour, Marie-therese; Haddada, Hedi; Lurquin,
     Christophe; Perricaudet, Michel; Uyttenhoveghesquiere, Catherine;
     Warnier, Guy
PA   Rhone-Poulenc Rorer S.A., Fr.; Ludwig Institute for Cancer Research;
     Boon, Falleur; Duffour, Marie-Therese; Haddada, Hedi; Lurquin,
     Christophe; Perricaudet, Michel; Uyttenhoveghesquiere, Catherine;
     Warnier, Guy
SO   PCT Int. Appl., 41 pp.
     CODEN: PIXXD2
PI   WO 9734009 A1  970918
DS   W:  AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GE, GH, HU, IL, IS,
         JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO,
         SG, SI, SK, TR, TT, UA, US, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD,
         RU, TJ, TM
     RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, DE, DK, ES, FI, FR, GA, GB,
         GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG
AI   WO 97-FR435  970312
PRAI FR 96-3207  960314
DT   Patent
LA   French
AB   A method for treating human tumors by gene therapy is disclosed.  In
     particular, defective recombinant viruses with a sequence coding for
     a human tumor-specific antigen, and the use thereof for treating or
     preventing human tumors, as well as producing specific cytotoxic
     T-cells (CTLs) in vitro or ex vivo, are disclosed.  Preferred
```

tumor-speciic antigens are Mage-1, Mage-3, Bage, Rage and Gage. Pharmaceutical compns. comprising said viruses, particularly in injectable form, are also disclosed.

IC ICM C12N015-86
ICS C07K014-705; A61K048-00; C12N005-10; A61K039-00; C12N015-12

CC 3-5 (Biochemical Genetics)
Section cross-reference(s): 1

IT DNA **sequences**
(of tumor antigen **Mage-1**, Mage-3 and P1A peptide-encoding DNA of mammal)

IT 197253-12-6
RL: BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)
(**Mage-1** peptide-encoding **sequence**; recombinant tumor-specific antigen-encoding adenoviral vectors for human tumor gene therapy)

L7 ANSWER 2 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1996:610305 HCAPLUS
DN 125:273611
TI Peptides derived from tumor rejection antigen precursor molecule MAGE-1 that bind MHC molecule HLA-C clone 10, and their uses
IN Van der Bruggen, Pierre; Boon-Falleur, Thierry
PA Ludwig Institute for Cancer Research, USA
SO U.S., 6 pp. Cont.-in-part of U.S. Ser. 8,446, abandoned.
CODEN: USXXAM
PI US 5558995 A 960924
AI US 94-195186 940214
PRAI US 93-8446 930122
DT Patent
LA English
AB The invention relates to the identification of complexes of HLA-C-clone 10 and MAGE-1 derived peptides on the surfaces of abnormal cells. These peptides presented by HLA-C-clone 10 are identical to those presented by HLA-A1. These peptides can therefore be used in the diagnosis of melanoma and in treatment by selection of pools of cytotoxic T-lymphocytes.
IC ICM C12N005-08
ICS C07K007-06; C07K007-08; G01N033-68
NCL 435007240
CC 15-2 (Immunochemistry)
IT 168836-14-4 170140-66-6 170173-06-5
RL: PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amino acid **sequence**; peptides derived from tumor rejection antigen precursor mol. **MAGE-1** that bind MHC mol. HLA-C clone 10, and their uses)

L7 ANSWER 3 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1996:494712 HCAPLUS
DN 125:165697
TI Monoclonal antibodies which bind to tumor rejection antigen precursor MAGE-1 and recombinant MAGE-1 oligopeptides
IN Chen, Yao-tseng; Stockert, Elisabeth; Chen, Yachi; Garin-chesa, Pilar; Rettig, Wolfgang J.; Van, Der Bruggen Pierre; Boon-falleur, Thierry; Old, Lloyd J.
PA Ludwig Institute for Cancer Research, USA; Cornell Research

Foundation, Inc.
SO U.S., 14 pp. Cont.-in-part of U.S. Ser. No. 37,280.
CODEN: USXXAM
PI US 5541104 A 960730
AI US 94-190411 940201
PRAI US 91-705702 910523
US 91-728838 910709
US 91-764365 910923
US 91-807043 911212
US 93-37230 930326
DT Patent
LA English
AB The invention relates to monoclonal antibodies which specifically bind to the tumor rejection antigen precursor mol. MAGE-1, hybridomas which produce these monoclonal antibodies, and their use. Also described is a recombinant form of MAGE-1, peptides which are useful as immunogens, and immunogenic compns. contg. the peptides and an adjuvant. In example, demonstrated were mol. cloning of recombinant MAGE-1 gene from human melanoma cell line MZ2-MEL 3.1, prepn. of peptides for raising monoclonal antibodies, binding of the monoclonal antibody 454 to testis lysate and MAGE-1-pos. melanomas, etc.
IC ICM C12P021-04
ICS C12P021-08; C12N005-00; C07K016-00
NCL 435240270
CC 15-3 (Immunochemistry)
Section cross-reference(s): 3
IT Deoxyribonucleic acid **sequences**
Protein **sequences**
Testis
(prepn. of monoclonal antibodies and recombinant peptides of tumor rejection antigen precursor **MAGE-1** as immunogen)
IT 180473-66-9
RL: PRP (Properties)
(nucleotide **sequence**; prepn. of monoclonal antibodies and recombinant peptides of tumor rejection antigen precursor **MAGE-1** as immunogen)

L7 ANSWER 4 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1996:22405 HCAPLUS
DN 124:53579
TI MAGE-1-specific precursor cytotoxic T-lymphocytes present among tumor-infiltrating lymphocytes from a patient with breast cancer: characterization and antigen-specific activation
AU Toso, John F.; Oei, Coreen; Oshidari, Farshid; Tartaglia, James; Paoletti, Enzo; Lyerly, H. Kim; Talib, Sohel; Weinhold, Kent J.
CS Departments of Pathology and Surgery, Duke University Medical Center, Durham, NC, 27710, USA
SO Cancer Res. (1996), 56(1), 16-20
CODEN: CNREA8; ISSN: 0008-5472
DT Journal
LA English
AB A potential target for development of tumor-specific immunotherapeutic strategies is the MAGE-1 gene. The authors have utilized a recently developed recombinant canarypox (ALVAC) virus vector contg. the MAGE-1 gene (vCP235) to activate CTLs from a

breast cancer patient bearing a MAGE-1+ tumor. Tumor-infiltrating lymphocytes (TILs) obtained from the tumor of a patient were stimulated in vitro with irradiated autologous peripheral blood mononuclear cells acutely infected with the vCP235 construct. These TILs preferentially expanded approx. 6-fold over a 16-day culture period and specifically recognized an allogeneic transformed B-cell line acutely infected with a vaccinia-MAGE-1 recombinant targeting vector (vP1188) in the context of HLA-A2 and/or B7. TCR V.beta. anal. of in vitro expanded T cells by a quant. multiprobe RNase protection assay revealed preferential expansion of TCR V.beta.6.3 and V.beta.6.4. In addn., homologous T-cell receptor .beta. CDR3 joining sequences were found in the in vitro stimulated cultures. These results suggest that tumor antigen-specific, MHC-restricted CTLs may be derived from precursor CTLs present in TILs obtained from patients with MAGE-1+ tumors by in vitro stimulation with recombinant avipox MAGE-1 virus-infected autologous cells. Collectively, these findings provide a rationale for tumor-assocd. antigen-based immunization as a means of activating precursor CTLs residing in patients with tumors expressing defined tumor-assocd. antigens such as MAGE-1.

CC 15-8 (Immunochemistry)
Section cross-reference(s): 3, 14

IT Protein **sequences**
(TCR .beta.-subunit VDJ CDR3 junction **sequences** in **MAGE-1**-specific precursor cytotoxic T-lymphocytes present among tumor-infiltrating lymphocytes from human with breast cancer)

IT Gene, animal
RL: BOC (Biological occurrence); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)
(TCR .beta.-subunit VDJ CDR3 junction **sequences** in **MAGE-1**-specific precursor cytotoxic T-lymphocytes present among tumor-infiltrating lymphocytes from human with breast cancer)

IT Antigen receptors
Receptors
RL: BOC (Biological occurrence); BIOL (Biological study); OCCU (Occurrence)
(TCR .alpha..beta. (T-cell antigen receptor .alpha..beta.), .beta.-subunit, genes for; TCR .beta.-subunit VDJ CDR3 junction **sequences** in **MAGE-1**-specific precursor cytotoxic T-lymphocytes present among tumor-infiltrating lymphocytes from human with breast cancer)

L7 ANSWER 5 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1995:874800 HCAPLUS
DN 123:283637
TI Monoclonal antibodies which bind to tumor rejection antigen precursor MAGE-1, recombinant MAGE-1, and MAGE-1-derived immunogenic peptides
IN Chen, Yao-Tseng; Stockert, Elisabeth; Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.; Van Der Bruggen, Pierre; Boon-Falleur, Thierry; Old, Lloyd J.
PA Ludwig Institute for Cancer Research, USA; Memorial Sloan-Kettering Cancer Center
SO PCT Int. Appl., 33 pp.
CODEN: PIXXD2

PI WO 9520974 A1 950810
DS W: AU, CA, CN, FI, JP, NO, NZ
RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE
AI WO 95-US95 950105
PRAI US 94-190411 940201
DT Patent
LA English
AB Monoclonal antibodies are provided which specifically bind to the tumor rejection antigen precursor mol. MAGE-1, as well as hybridomas which produce these monoclonal antibodies, and their use. Also described is a recombinant form of MAGE-1, peptides which are useful as immunogens, and immunogenic compns. contg. the peptides and an adjuvant.
IC ICM A61K038-04
ICS A61K038-10; A61K038-16; A61K039-395; A61K045-00; C07K014-46; C07K014-435; C07K016-30; C12N005-20; G01N033-53; G01N033-536
CC 15-3 (Immunochemistry)
Section cross-reference(s): 3
IT Protein **sequences**
(of tumor rejection antigen precursor **MAGE-1** from human)
IT Deoxyribonucleic acid **sequences**
(complementary, for tumor rejection antigen precursor **MAGE-1** from human)
IT 169440-60-2P
RL: BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(nucleotide **sequence**; monoclonal antibodies which bind to tumor rejection antigen precursor **MAGE-1**, recombinant **MAGE-1**, and **MAGE-1**-derived immunogenic peptides)

L7 ANSWER 6 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1995:826118 HCAPLUS
DN 123:225742
TI Multiple specificities in the repertoire of a melanoma patient's cytolytic T lymphocytes directed against tumor antigen MAGE-1.A1
AU Romero, Pedro; Pannetier, Christophe; Herman, Jean; Jongeneel, C. Victor; Cerottini, Jean-Charles; Coulie, Pierre G.
CS Ludwig Inst. for Cancer Res., Univ. of Lausanne, Epalinges, CH-1066, Switz.
SO J. Exp. Med. (1995), 182(4), 1019-28
CODEN: JEMEAV; ISSN: 0022-1007
DT Journal
LA English
AB Peptide MAGE-1.A1 is a nonamer derived from protein MAGE-1 that can assoc. with the HLA-A1 mol. It was shown previously to be recognized by an antitumor cytolytic T lymphocyte (CTL) clone derived from the blood of melanoma patient MZ2. We derived two other anti-MAGE-1.A1 CTL clones from different blood samples of the same patient and compared the fine specificity of recognition of the three CTL by testing them on variant MAGE-1.A1 peptides incorporating different amino acid substitutions. The epitopes recognized by the CTL proved to be different. While modifications of residues at positions 5, 6, or 7 in the antigenic peptide affected recognition by the three CTL, each of the modifications of residues at positions 1, 4, or 8 affected recognition by one CTL

only. The sequences of both the .alpha. and .beta. chains of the T cell antigen receptor of the three CTL were completely different. The results indicate a long-lasting diversity in terms of fine specificity and of T cell antigen receptor structure in the repertoire of antitumor CTL derived from the blood of a melanoma patient and directed against a defined tumor antigen.

CC 15-8 (Immunochemistry)

IT Protein **sequences**
(TCR receptor; multiple specificities in the repertoire of a melanoma human's cytolytic T lymphocytes directed against tumor antigen **MAGE-1**.A1)

L7 ANSWER 7 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1995:764993 HCAPLUS
DN 123:331445
TI Sequence analysis of the MAGE gene family encoding human tumor-rejection antigens
AU Imai, Yasuhisa; Shichijo, Shigeki; Yamada, Akira; Katayama, Takafumi; Yano, Hirohisa; Itoh, Kyogo
CS Department of Immunology, Kurume University School of Medicine, Kurume, 830, Japan
SO Gene (1995), 160(2), 287-90
CODEN: GENED6; ISSN: 0378-1119
DT Journal
LA English
AB The MAGE multigene family, which includes the MAGE-1 and -3 genes that encode tumor-rejection antigens on HLA-A1 recognized by cytotoxic T-lymphocytes (CTL), is preferentially expressed at the mRNA level on human malignant cells, but not on normal cells. However, little is known about the MAGE-4, MAGE-41, and MAGE-6 genes. In this study, 1040 bp (MAGE-1), 1061 bp (MAGE-3 and MAGE-6), and 1064 bp (MAGE-4 and MAGE-41) cDNA fragments, including the entire coding sequences (927-951 bp), were amplified using the reverse transcription-polymerase chain reaction (RT-PCR) method followed by nucleotide (nt) sequencing. One member had >80 or 66% homol. with the other members at the nt or deduced amino acid (aa) levels, resp. Higher homol. was found between MAGE-3 and -6 (98% at the nt level) and also between MAGE-4 and -41 (98%). The results of this investigation demonstrated high homol., as well as the clear differences between the members of the MAGE family at the coding sequence level.
CC 3-3 (Biochemical Genetics)
Section cross-reference(s): 6, 13
IT Gene, animal
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(**MAGE-1**; **sequence** anal. of the MAGE gene family encoding human tumor-rejection antigens)

L7 ANSWER 8 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1995:634555 HCAPLUS
DN 123:31227
TI Cloning and characterization of the complete MAGE-1 antigen gene and immunogenicity of its peptide fragments
IN Fikes, John D.; Livingston, Brian D.; Sette, Alessandro D.; Sidney, John C.
PA Cytel Corp., USA

SO PCT Int. Appl., 57 pp.
CODEN: PIXXD2
PI WO 9504542 A1 950216
DS W: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN
RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, DE, DK, ES, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG
AI WO 94-US8721 940802
PRAI US 93-103623 930806
DT Patent
LA English
AB The complete nucleotide and amino acid sequences of the human MAGE-1 antigen are provided. Peptides from residues of the C-terminal are used to define epitopes that stimulate HLA-restricted cytotoxic T lymphocyte activity against MAGE-1 antigens. The peptides are particularly useful in methods for stimulating the immune response of individuals against MAGE-1 antigens assocd. with melanomas.
IC ICM A61K038-00
ICS A61K039-00; C07H021-04; C12N001-19; C12N001-21; C12N005-10; C12N015-00; C12N015-70; C12N015-79; C12P021-06
CC 15-2 (Immunochemistry)
Section cross-reference(s): 1, 3
IT Deoxyribonucleic acid **sequences**
Genetic vectors
Immunostimulants
Melanoma
Molecular cloning
Protein **sequences**
(cloning and characterization of human **MAGE-1** antigen gene and immunogenicity of its peptide fragments)
IT 157298-43-6P 164251-72-3P
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(amino acid **sequence**; cloning and characterization of human **MAGE-1** antigen gene and immunogenicity of its peptide fragments)
IT 164108-94-5
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(nucleotide **sequence**; cloning and characterization of human **MAGE-1** antigen gene and immunogenicity of its peptide fragments)

L7 ANSWER 9 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1994:678846 HCAPLUS
DN 121:278846
TI Genes for tumor rejection antigens and the precursor MAGE-1 and their diagnostic and therapeutic uses
IN Boon, Thierry; van der Bruggen, Pierre; van den Eynde, Benoit; van Pel, Aline; de Plaen, Etienne; Lurquin, Christophe; Chomez, Patrick; Traversari, Catia
PA Ludwig Institute for Cancer Research, USA
SO U.S., 65 pp. Cont.-in-part of U.S. Ser. No. 764,364, abandoned.
CODEN: USXXAM

PI US 5342774 A 940830
AI US 91-807043 911212
PRAI US 91-705702 910523
US 91-728838 910709
US 91-764364 910923
DT Patent
LA English
AB Mouse genes for a group of tumor cell antigens that are recognized by cytotoxic T cells, leading to lysis of the tumor is cloned and expressed in animal cells for use in diagnostics and therapeutics. The genes were cloned by expression from a partial digest bank of tumor cell line DNA in cosmids using cytotoxic T-lymphocyte assays to detect expression. Presentation of the antigens was H-2 phenotype-dependent. Northern blots of tumors and normal tissue showed that the genes were expressed in a large array of tumors, but not in the normal tissues tested. CDNAs for several human mage precursor isoforms were cloned.
IC ICM C12P021-02
ICS C12P019-34; C12N015-00; C12N007-00; C12N005-00; C12N001-21; C12N001-16; C12N001-18; C07K003-00; C07H015-12
NCL 435240200
CC 15-2 (Immunochemistry)
Section cross-reference(s): 1, 9
IT 136361-96-1, Antigen (mouse clone C1A.3.1 gene P1A reduced)
RL: BOC (Biological occurrence); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); OCCU (Occurrence); USES (Uses)
(amino acid **sequence**; genes for tumor rejection antigens and the precursor **MAGE-1** and their diagnostic and therapeutic uses)
IT 140297-32-1, GenBank M36387 140767-62-0, GenBank M36386 146313-12-4, Deoxyribonucleic acid (human clone B3 gene **MAGE -1**) 146707-11-1, Deoxyribonucleic acid (mouse clone p815A gene p1A plus 5'- and 3'-flanking region fragment) 146707-15-5, Deoxyribonucleic acid (human melanoma gene MAGE-21 antigen fragment-specifying) 146707-16-6 146707-18-8, Deoxyribonucleic acid (human melanoma 1-121-gene MAGE-31 antigen-specifying plus 5'-flanking region fragment) 146707-29-1, Deoxyribonucleic acid (human melanoma gene MAGE-6 antigen fragment-specifying) 158968-97-9 158968-98-0 158968-99-1 158969-00-7 158969-01-8 158969-02-9
RL: BOC (Biological occurrence); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); OCCU (Occurrence); USES (Uses)
(nucleotide **sequence**; genes for tumor rejection antigens and the precursor **MAGE-1** and their diagnostic and therapeutic uses)

L7 ANSWER 10 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1994:555075 HCAPLUS
DN 121:155075
TI Cloning and analysis of MAGE-1-related genes
AU Ding, Min; Beck, Raymond J.; Keller, Christopher J.; Fenton, Robert G.
CS Clinical Research Branch, NCI, Frederick, MD, 21702, USA
SO Biochem. Biophys. Res. Commun. (1994), 202(1), 549-55
CODEN: BBRCA9; ISSN: 0006-291X
DT Journal
LA English

AB The spectrum of MAGE gene expression in the human melanoma cell line DM150 was examd. using reverse transcription polymerase chain reaction and cDNA cloning. Five full-length cDNAs were isolated from DM150 which were identified as MAGE-1, MAGE-3, MAGE-12, and 2 previously undescribed MAGE genes, MAGE-3b and MAGE-X2. DNA sequence anal. of the coding regions of the MAGE-3b and MAGE-X2 genes revealed 83 and 88% identity with MAGE-1, whereas MAGE-3b was 98% homologous with the full-length MAGE-3 clone. The predicted amino acid sequences of MAGE-X2 and MAGE-3b contain consensus HLA-A1 peptide binding motifs, suggesting that, like MAGE-1, they may code for tumor-assocd. antigens. In addn., a nonamer peptide encoded by both the MAGE-3 and MAGE-12 genes was shown by direct binding studies to contain an aggretope for HLA-A2.

CC 15-2 (Immunochemistry)
Section cross-reference(s): 3

IT Melanoma
(**MAGE-1**-related antigens of human, **sequence** and aggretope moieties of)

IT Gene, animal
RL: BIOL (Biological study)
(for melanoma-assocd. **MAGE-1**-related antigens, of human cell line DM150, **sequence** of)

IT Antigens
RL: BIOL (Biological study)
(melanoma-assocd. **MAGE-1**, amino acid **sequence** and aggretope moieties of, of human cell line DM150)

IT Protein **sequences**
(of melanoma-assocd. **MAGE-1**-related antigens, of human cell line DM150)

IT Deoxyribonucleic acid **sequences**
(complementary, for melanoma-assocd. **MAGE-1** -related antigens, of human cell line DM150)

IT 157298-41-4, Melanoma-assocd. antigen MAGE-X2 (human cell line DM150) 157298-42-5, Melanoma-assocd. antigen MAGE-3b (human cell line DM150) 157298-43-6, Melanoma-assocd. antigen **MAGE-1** (human cell line DM150) 157298-44-7, Melanoma-assocd. antigen MAGE-12f (human cell line DM150)
RL: BIOL (Biological study)
(amino acid **sequence** and aggretope moieties of)

L7 ANSWER 11 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1993:167452 HCAPLUS
DN 118:167452
TI Cloning of genes for tumor rejection antigen precursors and their uses
IN Boon, Thierry; Van der Bruggen, Pierre; Van den Eynde, Benoit; Van Pel, Aline; De Plaen, Etienne; Lurquin, Christophe; Chomez, Patrick; Traversari, Catia
PA Ludwig Institute for Cancer Research, USA
SO PCT Int. Appl., 143 pp.
CODEN: PIXXD2
PI WO 9220356 A1 921126
DS W: AU, BB, BG, BR, CA, CS, FI, HU, JP, KP, KR, LK, MG, MW, NO, PL, RO, RU, SD, US
RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, DE, DK, ES, FR, GA, GB, GR, IT, LU, MC, ML, MR, NL, SE, SN, TD, TG

AI WO 92-US4354 920522
PRAI US 91-705702 910523
US 91-728838 910709
US 91-764364 910923
US 91-807043 911212
DT Patent
LA English
AB The genes or cDNA for tumor rejection antigen (TRA) precursors, e.g. the precursors for MAGE melanoma antigen, the P1A mastocytoma antigen, and the antigen F of human, and the murine counterpart of MAGE, smage, are cloned. The coding sequences of these TRA can be used for prepn. of vaccines by expression of the sequences alone or together with the gene for a cytokine, e.g., interleukin (IL)-2 or IL-4. They can also be expressed with a gene for an MHC or HLA antigen which presents the tumor rejection antigen derived from the precursor to the cytotoxic T cells. Expression of the TRA by tumor cells can lead to cell lysis mediated by the cytotoxic T cells that recognize the antigens. The TRA may be used for prepn. of pharmaceuticals, antibodies, and diagnostics for clin. applications.
IC A61K035-14; A61K039-00; A61K037-22; C07K003-00; C07K015-00; C07K017-00; C12Q001-68; C12Q001-00; C12Q015-00
CC 15-2 (Immunochemistry)
Section cross-reference(s): 1, 3, 9
IT 140101-03-7, Deoxyribonucleic acid (human clone B3 gene **MAGE -1** plus 5'- and 3'-flanking region fragment) 146707-11-1
146707-12-2 146707-13-3 146707-15-5 146707-16-6 146707-18-8
146707-20-2 146707-22-4 146707-24-6 146707-25-7 146707-27-9
146707-29-1 146707-30-4 146707-32-6 146707-34-8 146707-36-0
146707-38-2 146707-40-6 146707-42-8
RL: BIOL (Biological study); PRP (Properties)
(nucleotide **sequence** and cloning of)
IT 136362-51-1, Deoxyribonucleic acid (mouse clone C1A.3.1 gene P1A coding region) 146313-12-4, Deoxyribonucleic acid (human clone B3 gene **MAGE-1**) 146707-14-4 146707-17-7
146707-19-9 146707-21-3 146707-23-5 146707-26-8 146707-28-0
146707-31-5 146707-33-7 146707-35-9 146707-37-1 146707-39-3
146707-41-7
RL: BIOL (Biological study); PRP (Properties)
(nucleotide **sequence** and cloning of, complete)

L7 ANSWER 12 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1993:121916 HCAPLUS
DN 118:121916
TI A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma
AU Van der Bruggen, P.; Traversari, C.; Chomez, P.; Lurquin, C.; De Plaen, E.; Van den Eynde, B.; Knuth, A.; Boon, T.
CS Ludwig Inst. Cancer Res., Brussels, B-1200, Belg.
SO Science (Washington, D. C., 1883-) (1991), 254(5038), 1643-7
CODEN: SCIEAS; ISSN: 0036-8075
DT Journal
LA English
AB Many human melanoma tumors express antigens that are recognized in vitro by cytolytic T lymphocytes (CTLs) derived from the tumor-bearing patient. A gene was identified that directed the expression of antigen MZ2-E on a human melanoma cell line. This gene shows no similarity to known sequences and belongs to a family

of at least three genes. It is expressed by the original melanoma cells, other melanoma cell lines, and by some tumor cells of other histol. types. No expression was obsd. in a panel of normal tissues. Antigen MZ2-E appears to be presented by HLA-A1; anti-MZ2-E CTLs of the original patient recognized two melanoma cell lines of other HLA-A1 patients that expressed the gene. Thus, precisely targeted immunotherapy directed against antigen MZ2-E could be provided to individuals identified by HLA typing and anal. of the RNA of a small tumor sample.

CC 14-1 (Mammalian Pathological Biochemistry)
Section cross-reference(s): 3, 15

IT 146313-13-5, Antigen (human clone B3 gene **MAGE-1** reduced)
RL: PRP (Properties)
(amino acid **sequence** of, complete)

IT 140101-03-7, Deoxyribonucleic acid (human clone B3 gene **MAGE -1** plus 5'- and 3'-flanking region fragment)
RL: PRP (Properties)
(nucleotide **sequence** of)

IT 146313-12-4, Deoxyribonucleic acid (human clone B3 gene **MAGE -1**)
RL: PRP (Properties)
(nucleotide **sequence** of, complete)

L7 ANSWER 13 OF 13 HCAPLUS COPYRIGHT 1998 ACS
AN 1992:631775 HCAPLUS
DN 117:231775
TI A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic T lymphocytes directed against tumor antigen MZ2-E
AU Traversari, Catia; Van der Bruggen, Pierre; Luescher, Immanuel F.; Lurquin, Christophe; Chomez, Patrick; Van Pel, Aline; De Plaen, Etienne; Amar-Costesec, Alain; Boon, Thierry
CS Brussels Branch, Ludwig Inst. Cancer Res., Brussels, B-1200, Belg.
SO J. Exp. Med. (1992), 176(5), 1453-7
CODEN: JEMEAV; ISSN: 0022-1007
DT Journal
LA English
AB The authors reported the identification of human gene MAGE-1, which directs the expression of an antigen recognized on a melanoma by autologous cytolytic T lymphocytes (CTL). CTL directed against this antigen, which was named MZ2-E, recognize a nonapeptide encoded by the third exon of gene MAGE-1. The CTL also recognize this peptide when it is presented by mouse cells transfected with an HLA-A1 gene, confirming the assocn. of antigen MZ2-E with the HLA-A1 mol. Other members of the MAGE gene family do not code for the same peptide, suggesting that only MAGE-1 produces the antigen recognized by the anti-MZ2-E CTL. The results open the possibility of immunizing HLA-A1 patients whose tumor expresses MAGE-1 either with the antigenic peptide or with autologous antigen-presenting cells pulsed with the peptide.

CC 15-2 (Immunochemistry)

IT Protein **sequences**
(of gene **MAGE-1**-encoded antigen, of humans)

ALL IS NOT A RECOGNIZED COMMAND

=> d all l16

L16 ANSWER 1 OF 1 HCAPLUS COPYRIGHT 1998 ACS
AN 1990:53087 HCAPLUS
DN 112:53087
TI Synthesis of vitamin K-dependent **proteins** by cultured human **tumor** cells
AU Al-Mondhiry, Hamid; Wallin, Reidar
CS Coll. Med., Pennsylvania State Univ., Hershey, PA, 17033, USA
SO Thromb. Haemostasis (1989), 62(2), 661-6
CODEN: THHADQ; ISSN: 0340-6245
DT Journal
LA English
CC 14-1 (Mammalian Pathological Biochemistry)
Section cross-reference(s): 13
AB The observation that warfarin inhibits the growth and metastasis of certain types of clin. and exptl. tumors suggests a role for vitamin K in tumor biol. The synthesis of vitamin K-dependent proteins was investigated in four malignant (lung epidermoid carcinoma, melanoma, colon adenocarcinoma, and breast adenocarcinoma) and three normal (colon epithelium, breast epithelium, and fibroblasts) cell lines of human origin grown in tissue cultures. The results show the following: 1) vitamin K-dependent carboxylase activity is present in all of the malignant and normal cell lines studied; 2) the malignant, as well as normal, cell lines synthesize a family of vitamin K-dependent proteins, and microsomal precursors of these proteins with apparent mol. masses of 74, 62, and 34 kDa are common to all malignant and normal cell lines, whereas precursors of higher and lower mol. mass seem to be synthesized by some but not all tumor cell lines; and 3) the 74-kDa precursor synthesized by colon carcinoma and breast carcinoma was pos. identified as a precursor of protein S.
ST vitamin K **protein** formation **tumor** cell
IT **Melanoma**
Neoplasm, metabolism
(vitamin K-dependent **protein** formation by, of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)
IT Fibroblast
(vitamin K-dependent protein formation by, of humanso)
IT Protein formation
(vitamin K-dependent, by neoplastic and normal cells of humans)
IT Carcinoma
(adeno-, vitamin K-dependent protein formation by, of intestine and mammary gland of humans)
IT Intestine, neoplasm
(colon, adenocarcinoma, vitamin K-dependent **protein** formation by, of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)
IT Intestine, metabolism
(colon, epithelium, vitamin K-dependent protein formation by, of humans)
IT Blood coagulation
(disorder, hypercoagulability, vitamin K-dependent protein formation by human neoplastic cells in relation to)

IT Carcinoma
(epidermoid, vitamin K-dependent protein formation by, of lung of human)
IT Lung, neoplasm
(epidermoid carcinoma, vitamin K-dependent **protein** formation by, of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)
IT Mammary gland
(epithelium, vitamin K-dependent protein formation by, of humans)
IT Mammary gland
(neoplasm, adenocarcinoma, vitamin K-dependent **protein** formation by, of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)
IT Blood-coagulation factors
RL: BIOL (Biological study)
(protein S, precursor of, formation of, by colon carcinoma and breast carcinoma cells, of humans)
IT Proteins, specific or class
RL: BIOL (Biological study)
(.gamma.-carboxyglutamic acid-contg., 23,000-**mol**.-**wt**., formation of vitamin K-dependent, by neoplastic cells of humans)
IT Proteins, specific or class
RL: BIOL (Biological study)
(.gamma.-carboxyglutamic acid-contg., 30,000-**mol**.-**wt**., formation of vitamin K-dependent, by neoplastic cells of humans)
IT **Proteins**, specific or class
RL: BIOL (Biological study)
(.gamma.-carboxyglutamic acid-contg., **34**,000-**mol**.-**wt**., formation of vitamin K-dependent, by neoplastic and normal cells of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)
IT **Proteins**, specific or class
RL: BIOL (Biological study)
(.gamma.-carboxyglutamic acid-contg., 62,000-**mol**.-**wt**., formation of vitamin K-dependent, by neoplastic and normal cells of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)
IT **Proteins**, specific or class
RL: BIOL (Biological study)
(.gamma.-carboxyglutamic acid-contg., 74,000-**mol**.-**wt**., formation of vitamin K-dependent, by neoplastic and normal cells of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)
IT Proteins, specific or class
RL: BIOL (Biological study)
(.gamma.-carboxyglutamic acid-contg., 89,000-**mol**.-**wt**., formation of vitamin K-dependent, by neoplastic cells of humans)
IT 12001-79-5, Vitamin K
RL: BIOL (Biological study)
(protein formation dependent on, of normal and neoplastic human cells)

```
<------User Break------>

=> fil wpids

FILE 'WPIDS' ENTERED AT 09:22:01 ON 07 APR 1998
COPYRIGHT (C) 1998 DERWENT INFORMATION LTD

FILE LAST UPDATED: 06 APR 1998              <19980406/UP>
>>>UPDATE WEEKS:
MOST RECENT DERWENT WEEK            199814   <199814/DW>
DERWENT WEEK FOR CHEMICAL CODING:   199809
DERWENT WEEK FOR POLYMER INDEXING:  199811
DERWENT WORLD PATENTS INDEX SUBSCRIBER FILE, COVERS 1963 TO DATE
>>> D COST AND SET NOTICE DO NOT REFLECT SUBSCRIBER DISCOUNTS -
                                    SEE HELP COST FOR DETAILS <<<
>>> MEXICO NOW COVERED  - SEE NEWS <<<

=> d his


     (FILE 'WPIDS' ENTERED AT 09:17:35 ON 07 APR 1998)
                DEL HIS Y
L1            9 S MAGE 1
L2            2 S 11540
L3            0 S L2 AND (ANTIBOD? OR MONOCLONAL?)
L4         1459 S (TUMOR# OR TUMOUR#) (6A) (PROTEIN# OR ANTIGEN#)
L5         1403 S MELANOMA
L6           96 S L4 AND L5
L7            0 S L6 AND (46 OR 34)
L8            8 S L6 AND (KD OR KDS OR KILODALTON# OR KILO DALTON# )

     FILE 'WPIDS' ENTERED AT 09:21:18 ON 07 APR 1998

     FILE 'WPIDS' ENTERED AT 09:22:01 ON 07 APR 1998

=> d .wp l1 1-9; d .wp l8 1-8


L1   ANSWER 1 OF 9  WPIDS   COPYRIGHT 1998 DERWENT INFORMATION LTD
AN   98-086733 [08]   WPIDS
DNC  C98-029335
TI   Inducing anti-tumour cytotoxic T lymphocytes  - by cross-priming
     with artificial antigen, then immunisation with tumour cells
     modified in vitro to express the same antigen, does not require
     characterisation of tumour-specific antigens.
DC   B04 D16
IN   FALO, L D; ROCK, K L
PA   (DAND) DANA FARBER CANCER INST INC; (UYPI-N) UNIV PITTSBURGH
CYC  77
PI   WO 9800163  A1 980108 (9808)* EN   44 pp
        RW: AT BE CH DE DK EA ES FI FR GB GH GR IE IT KE LS LU MC MW NL
            OA PT SD SE SZ UG ZW
         W: AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI
            GB GE GH HU IL IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD
            MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR
            TT UA UG UZ VN YU ZW
```

ADT WO 9800163 A1 WO 97-US10195 970618
PRAI US 96-675332 960628
AB WO 9800163 A UPAB: 980223
Anti-tumour immunisation in a mammal comprises: (a) immunising the host with an artificial target antigen (A) to produce a cytotoxic T lymphocyte (CTL) response; (b) resecting tumour from the host; (c) culturing tumour cells in vitro; (d) engineering these cells to include (A) so that presentation of (A) on the cell surface is promoted, and (e) inactivating the cultured cells and returning them to the host. In a modification, step (a) is omitted and the engineered cells are returned to the host simultaneously with immunisation by (A).

In step (a), (A) is administered as a particulate complex, e.g. by injection or particle bombardment. In step (d), cells are transfected with nucleic acid expressing (A), or (A) is introduced by peptide pulsing. Specified (A) are (i) tumour antigens such as Melan-A, p53, carcinoembryonal antigen, gp100, **MAGE-1** or -2; (ii) a viral antigen such as HIV gp120 or gp100, influenza virus nucleoprotein or hepatitis B surface antigen or (iii) chicken ovalbumin (OVA, especially preferred) or keyhole limpet haemocyanin.

USE - The method is used for treatment or prevention of tumours, specifically in humans. A similar method is used (not claimed) for treatment of viral infections, e.g. with human immunodeficiency virus (HIV).

ADVANTAGE - The method of cross-priming results in a CTL response to many, undefined, MHC Class I-restricted tumour antigens expressed on the surface of unmodified tumour cells. It can be applied to a wide range of tumours and does not require isolation and characterisation of particular tumour (or viral) antigens.
Dwg.5/8

L1 ANSWER 2 OF 9 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 97-212603 [19] WPIDS
DNC C97-068627
TI Use of DNA coated particles - for therapeutic or prophylactic immunisation of mammals, partic. for neoplastic cells or virus infected cells.
DC B04 D16
IN FALO, L D; ROCK, K L
PA (DAND) DANA FARBER CANCER INST INC; (UYPI-N) UNIV PITTSBURGH
CYC 69
PI WO 9711605 A1 970403 (9719)* EN 46 pp
RW: AT BE CH DE DK EA ES FI FR GB GR IE IT KE LS LU MC MW NL OA PT SD SE SZ UG
W: AL AM AT AU AZ BB BG BR BY CA CH CN CZ DE DK EE ES FI GB GE HU IS JP KE KG KP KR KZ LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK TJ TM TR TT UA UG UZ VN
AU 9672515 A 970417 (9732)
ADT WO 9711605 A1 WO 96-US15728 960927; AU 9672515 A AU 96-72515 960927
FDT AU 9672515 A Based on WO 9711605
PRAI US 95-535556 950928
AB WO 9711605 A UPAB: 970512
In vivo method (A) of therapeutic or prophylactic genetic immunisation of a mammalian host comprises:

(a) generating a DNA fragment which expresses an antigenic protein or fragment,

(b) distributing the DNA fragment on a particle surface, resulting in a particulate polynucleotide (PP),
(c) inoculating the mammalian host with the PP and
(d) delivering the PP to the cytoplasm of a target cell within the mammalian host, such that the expressed antigenic protein or fragment is presented to the membrane surface of the target cell through the major histocompatibility complex (MHC) class I pathway.
Also claimed are:
(1) an ex vivo method (B) of therapeutic or prophylactic genetic immunisation of a mammalian host, comprising:
(a) generating a DNA fragment which expresses an antigenic protein or fragment,
(b) distributing the DNA fragment on a particle surface, resulting in a PP,
(c) delivering the PP to the cytoplasm of a target cell of a mammalian host in vitro, such that the expressed antigenic protein or fragment is presented on the membrane surface of the target cell through the MHC class I pathway and
(d) inoculating the mammalian host with the target cell by direct injection, and
(3) an ex vivo method (C) of therapeutic or prophylactic genetic immunisation of a mammalian host, comprising:
(a) generating a DNA fragment which expresses a molecule which enhances the antigen presentation function of an antigen presenting cell (APC),
(b) distributing the DNA fragment on a particle surface, resulting in a PP,
(c) delivering the PP to the cytoplasm of a target cell of a mammalian host in vitro, such that the expressed antigen presentation enhancing protein is expressed in a biologically significant form and at biologically significant levels and
(d) inoculating the mammalian host with the target cell by direct injection.
USE - The methods are used for producing antigen-specific immune responses to antigens such as tumour rejection antigens, e.g. **MAGE-1**, MAGE-3, Melan+A, gp100, p53, CEA or HER2/neu or viral antigens, eg. HIV gp120, HIV gp160, influenza virus nucleoprotein or hepatitis B surface antigen. They can be used for therapeutic or prophylactic treatment (all claimed).
ADVANTAGE - The PP can target delivery of antigens to APCs to increase the efficiency of genetic immunisation while reducing unwanted deleterious effects. Proper presentation of the antigenic peptides through the MHC class I pathway stimulates cytotoxic T lymphocyte (CTL) production and in turns promoters destruction of cells such as neoplastic cells or virally infected cells.
Dwg.0/6

L1 ANSWER 3 OF 9 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 97-108910 [10] WPIDS
DNN N97-090092 DNC C97-034778
TI Prepn. of photo-reactive peptide derivs. - by substitution. with photoreactive amino acid then radio-iodination, giving prods. used to assess ability of peptide to bind to specific major histocompatibility complex mols..
DC B04 D16 K08 S03
IN ANJUERE, F; CERROTINI, J; LAYERE, A; LEUSCHER, I; ROMERO, P
PA (LUDW-N) LUDWIG INST CANCER RES

CYC 23
PI WO 9702282 A1 970123 (9710)* EN 36 pp
RW: AT BE CH DE DK ES FI FR GB GR IE IT LU MC NL PT SE
W: AU CA CN FI JP NO NZ
AU 9665418 A 970205 (9721)
ADT WO 9702282 A1 WO 96-US10869 960625; AU 9665418 A AU 96-65418 960625
FDT AU 9665418 A Based on WO 9702282
PRAI US 95-498461 950705
AB WO 9702282 A UPAB: 970326
Prepn. of synthetic photoreactive peptide derivs. (I) comprises:
(a) linear synthesis of a peptide;
(b) substitution. of an amino acid (aa) in the peptide with a photoreactive aa at a position that does not alter its binding ability, and
(c) specific radio-iodination of the photoreactive aa.
Also new are peptides of formula (Ia) Q-X-A-Tyr (Ia)
Q = photoreactive gp., specifically iodinated 2,3-(4-azido-salicyloyl)-diaminopropionic acid (AzDAP);
X = Ala or Val;
A = any 6 aa.
USE - The method is esp. applied to the melanoma-derived MAGE peptides, esp. **MAGE-1**, -3, -4a,-6 or -12, to give cpds. used to assess ability of a peptide to bind to specific MHC mols. Peptides which bind specifically, are used to screen either MHC mols. for cross-reactivity or peptides for their ability to inhibit the photoaffinity labelling reaction. A partic. application is identification of MHC able to present a given cytotoxic T cell epitope.
ADVANTAGE - The method provides a radio-iodination yield of over 90% without significant formation of di-iodinated prod., so that HPLC purificn. of the prod. is not required and reverse-phase purificn. provides complete sepn. of the prod. from non-iodinated precursors. Photoaffinity labelling is very specific, so does require isolation of MHC from cells, making possible rapid testing of panels of peptides or cell lines expressing different MHC.
Dwg.0/0

L1 ANSWER 4 OF 9 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 95-320586 [41] WPIDS
CR 92-415460 [50]; 94-100844 [12]; 94-333192 [41]; 95-283606 [37]
DNC C95-142446
TI Determn. of cancerous condition(s) - using a nucleic acid as a primer to determine expression of a MAGE tumour rejection antigen precursor.
DC B04 D16
IN BOON-FALLEUR, T; BRASSEUR, F; CHOMEZ, P; DE PLAEN, E; DE SMET, C; GAUGLER, B; LETHE, B; MARCHAND, M; PATARD, J; SZIKORA, J; VAN DEN EYNDE, B; VAN DER BRUGGEN, P; WEYNANTS, P; BASSEUR, F; DEPLAEN, E
PA (LUDW-N) LUDWIG INST CANCER RES
CYC 24
PI WO 9523874 A1 950908 (9541)* EN 122 pp
RW: AT BE CH DE DK ES FR GB GR IE IT LU MC NL PT SE
W: AU CA FI JP KR NO NZ
AU 9519682 A 950918 (9551)
US 5512437 A 960430 (9623) 5 pp
US 5512444 A 960430 (9623) 10 pp
FI 9603393 A 960830 (9646)

NO 9603589 A 961031 (9702)
US 5612201 A 970318 (9717) 72 pp
JP 09509832 W 971007 (9750) 132 pp
ADT WO 9523874 A1 WO 95-US2203 950223; AU 9519682 A AU 95-19682 950223; US 5512437 A US 94-204727 940301; US 5512444 A CIP of US 94-204727 940301, US 94-346774 941130; FI 9603393 A WO 95-US2203 950223, FI 96-3393 960830; NO 9603589 A WO 95-US2203 950223, NO 96-3589 960828; US 5612201 A CIP of US 91-705702 910523, CIP of US 91-728838 910709, CIP of US 91-764364 910923, CIP of US 91-807043 911212, CIP of WO 92-US4354 920522, CIP of US 93-37230 930326, US 94-299849 940901; JP 09509832 W JP 95-522934 950223, WO 95-US2203 950223
FDT AU 9519682 A Based on WO 9523874; US 5612201 A CIP of US 5327252, CIP of US 5342774; JP 09509832 W Based on WO 9523874
PRAI US 94-346774 941130; US 94-204727 940301; US 94-209172 940310; US 94-299849 940901; US 91-705702 910523; US 91-728838 910709; US 91-764364 910923; US 91-807043 911212; WO 92-US4354 920522; US 93-37230 930326
AB WO 9523874 A UPAB: 980126
Isolated nucleic acid (NA) useful as a primer in determining the expression of a member of the melanoma antigen (MAGE) group of tumour rejection antigen precursors selected from one of 20 given sequences. Also claimed is a kit for determining the expression of a MAGE tumour rejection antigen precursor comprising at least 1 specific pair of the above NA sequences.
USE - The NA is useful in cancer determination assays, esp. melanoma, bead (sic) or neck squamous cell carcinoma, prostate carcinoma or bladder tumour (claimed).
Dwg.0/20

L1 ANSWER 5 OF 9 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 95-292948 [38] WPIDS
CR 94-263764 [32]; 95-051741 [07]
DNC C95-131893
TI Identification of cells presenting HLA-C-clone 10 or **MAGE-1** derived peptide - allows diagnosis and treatment of individuals with cellular abnormalities, e.g. melanoma, also HLA-Cw 1601 derived peptide(s).
DC B04 D16
IN BOEL, P; BOON-FALLEUR, T; COULIE, P; SZIKORA, J; VAN, DER BRUGGEN P; WILDMANN, C; WILDMAN, C
PA (LUDW-N) LUDWIG INST CANCER RES
CYC 26
PI WO 9521630 A1 950817 (9538)* EN 26 pp
RW: AT BE CH DE DK ES FR GB GR IE IT LU MC NL PT SE
W: AU CA CN FI JP NO NZ
AU 9520899 A 950829 (9548)
ZA 9501177 A 960424 (9622) 30 pp
FI 9603170 A 960813 (9644)
NO 9603347 A 960814 (9644)
US 5558995 A 960924 (9644) 6 pp
EP 789591 A1 970820 (9738) EN
R: AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE
ADT WO 9521630 A1 WO 95-US1446 950126; AU 9520899 A AU 95-20899 950126; ZA 9501177 A ZA 95-1177 950214; FI 9603170 A WO 95-US1446 950126, FI 96-3170 960813; NO 9603347 A WO 95-US1446 950126, NO 96-3347 960812; US 5558995 A CIP of US 93-8446 930122, US 94-195186 940214; EP 789591 A1 EP 95-913476 950126, WO 95-US1446 950126

FDT AU 9520899 A Based on WO 9521630; EP 789591 A1 Based on WO 9521630
PRAI US 94-292492 940818; US 94-195186 940214; US 94-196630 940215; US 93-8446 930122
AB WO 9521630 A UPAB: 961111
A novel method for identifying a candidate for treatment with a therapeutic agent specific for complexes of HLA-C-clone 10 and the peptide (I): Ser-Ala-Tyr-Gly-Glu-Pro-Arg-Lys-Leu (I) comprises: (i) contacting an abnormal cell sample from a subject with a cytolytic T cell specific for the complexes; and (ii) determining lysis of at least part of the abnormal cell sample as an indication of a candidate for the treatment.
USE - The method can be used to treat a subject with a cellular abnormality eg melanoma or to identify an abnormal cell which presents a complex of HLA-C-clone 10 and (I) on its surface. The isolated peptides are useful in diagnosing cancer. The isolated nucleic acid mols are useful as probes for the determination of expression of HLA-Cw*1601. HLA typing is important in tissue typing for transplantation.
Dwg.0/2

L1 ANSWER 6 OF 9 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 95-283606 [37] WPIDS
CR 92-415460 [50]; 94-100844 [12]; 94-333192 [41]; 95-320586 [41]
DNN N95-215847 DNC C95-127959
TI New monoclonal antibody binding specifically to **MAGE-1** - useful for diagnosis and monitoring of cancer, also new hybridomas, recombinant **MAGE-1** and immunogenic peptide(s).
DC B04 D16 S03
IN BOON-FALLEUR, T; CHEN, Y; GARIN-CHESA, P; OLD, L J; RETTIG, W J; STOCKERT, E; VAN DER BRUGGEN, P; OLD, L
PA (LUDW-N) LUDWIG INST CANCER RES; (SLOK) SLOAN KETTERING INST CANCER RES; (CORR) CORNELL RES FOUND INC; (SLOK) MEMORIAL SLOAN-KETTERING CANCER CENT
CYC 26
PI WO 9520974 A1 950810 (9537)* EN 33 pp
RW: AT BE CH DE DK ES FR GB GR IE IT LU MC NL PT SE
W: AU CA CN FI JP NO NZ
AU 9515979 A 950821 (9547)
ZA 9500786 A 951227 (9605) 29 pp
US 5541104 A 960730 (9636) 14 pp
FI 9603033 A 960731 (9642)
NO 9603120 A 960930 (9649)
EP 752876 A1 970115 (9708) EN
R: AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE
JP 09511389 W 971118 (9805) 36 pp
AU 686314 B 980205 (9813)
ADT WO 9520974 A1 WO 95-US95 950105; AU 9515979 A AU 95-15979 950105; ZA 9500786 A ZA 95-786 950201; US 5541104 A CIP of US 91-705702 910523, CIP of US 91-728838 910709, CIP of US 91-764365 910923, CIP of US 91-807043 911212, CIP of WO 92-US4354 920522, CIP of US 93-37230 930326, US 94-190411 940201; FI 9603033 A WO 95-US95 950105, FI 96-3033 960731; NO 9603120 A WO 95-US95 950105, NO 96-3120 960726; EP 752876 A1 EP 95-907978 950105, WO 95-US95 950105; JP 09511389 W JP 95-520611 950105, WO 95-US95 950105; AU 686314 B AU 95-15979 950105
FDT AU 9515979 A Based on WO 9520974; US 5541104 A CIP of US 5342774; EP

752876 A1 Based on WO 9520974; JP 09511389 W Based on WO 9520974; AU 686314 B Previous Publ. AU 9515979, Based on WO 9520974

PRAI US 94-190411 940201; US 91-705702 910523; US 91-728838 910709; US 91-764365 910923; US 91-807043 911212; WO 92-US4354 920522; US 93-37230 930326

AB WO 9520974 A UPAB: 980126

Monoclonal antibody (MAb) that binds specifically to tumour rejection antigen precursor **MAGE-1** is new. Also new are (1) hybridomas that produce MAb; (2) isolated **MAGE -1** deriv. (A) of mol. wt. 20-22 kD; (3) isolated protein (A1) comprising amino acids 57-219 encoded by nucleotides 3931-4761 of the **MAGE-1** gene (sequence of 5674 bp reproduced); and (4) isolated peptides of formulae: Ile Asn Phe Thr Arg Gln Arg Gln Pro Ser Glu Gly Ser Ser (2) Leu Phe Arg Ala Val Ile Thr Lys Lys Val Ala Asp (3) Asp Val Lys Glu Ala Asp Pro Thr Gly His Ser Tyr (4).

USE - MAb, opt. labelled or immobilised, is used to detect **MAGE-1** in samples (e.g. cell lysates) by standard immunoassay methods, e.g. for diagnosis and monitoring of cancer.(A), (A1) and peptides (2)-(4) are useful as immunogens for prodn. of MAb and antisera.

ADVANTAGE - MAb is specific for **MAGE-1**, having no reactivity for MAGE-2 or 3.

Dwg.0/4

L1 ANSWER 7 OF 9 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD

AN 95-090681 [12] WPIDS

DNC C95-041028

TI Human melanoma antigen, **MAGE-1**, peptide(s) - useful for stimulating immune response against melanoma.

DC B04 D16

IN FIKES, J D; LIVINGSTON, B D; SETTE, A D; SIDNEY, J C

PA (CYTE-N) CYTEL CORP

CYC 57

PI WO 9504542 A1 950216 (9512)* EN 59 pp
RW: AT BE CH DE DK ES FR GB GR IE IT KE LU MC MW NL OA PT SD SE
W: AM AT AU BB BG BR BY CA CH CN CZ DE DK ES FI GB GE HU JP KE KG KP KR KZ LK LT LU LV MD MG MN MW NL NO NZ PL PT RO RU SD SE SI SK TJ TT UA US UZ VN
AU 9475534 A 950228 (9524)
EP 721341 A1 960717 (9633) EN
R: AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE
JP 09502086 W 970304 (9719) 60 pp
NZ 271774 A 980226 (9813)

ADT WO 9504542 A1 WO 94-US8721 940802; AU 9475534 A AU 94-75534 940802; EP 721341 A1 EP 94-925722 940802, WO 94-US8721 940802; JP 09502086 W WO 94-US8721 940802, JP 95-506486 940802; NZ 271774 A NZ 94-271774 940802, WO 94-US8721 940802

FDT AU 9475534 A Based on WO 9504542; EP 721341 A1 Based on WO 9504542; JP 09502086 W Based on WO 9504542; NZ 271774 A Based on WO 9504542

PRAI US 93-103623 930806

AB WO 9504542 A UPAB: 950328

Immunogenic peptides, comprising ca. 9-10 residues from the C-terminal of melanoma antigen protein **MAGE-1** (58 residue sequence given in specification), are new. Also claimed are: (1) isolated DNA encoding the peptides; (2) vectors comprising the DNA; and (3) host cells transformed with the vectors.

USE - The peptides are useful for defining epitopes stimulating HLA-restricted cytotoxic T-lymphocyte activity against **MAGE -1** antigens (Ags). They are particularly useful for stimulating the immune response of individuals against **MAGE -1** Ags associated with melanomas. Vaccine compositions containing these peptides can be administered to a patient susceptible to MAGE-associated tumours, eg. melanomas. Dwg.0/3

L1 ANSWER 8 OF 9 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 94-263764 [32] WPIDS
CR 95-051741 [07]; 95-292948 [38]
DNC C94-120655
TI Treating and detecting, cellular abnormality with specific cytolytic T cells - recognising particular HLA-MAGE - cell surface complex, or with agents which provoke such cells, esp. for melanoma.
DC B04 D16 P31
IN BOON-FALLEUR, T; VAN, DER BRUGGEN P
PA (LUDW-N) LUDWIG INST CANCER RES
CYC 26
PI WO 9416713 A1 940804 (9432)* EN 17 pp
RW: BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE
W: AU CA FI JP NO NZ
AU 9460913 A 940815 (9444)
ZA 9400407 A 941026 (9444) 18 pp
FI 9503534 A 950721 (9542)
NO 9502797 A 950920 (9547)
EP 680330 A1 951108 (9549) EN
R: AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE
JP 08506015 W 960702 (9650) 16 pp
CN 1096584 A 941221 (9718)
US 5629166 A 970513 (9725) 6 pp
AU 680236 B 970724 (9737)
ADT WO 9416713 A1 WO 94-US688 940118; AU 9460913 A AU 94-60913 940118; ZA 9400407 A ZA 94-407 940120; FI 9503534 A WO 94-US688 940118, FI 95-3534 950721; NO 9502797 A WO 94-US688 940118, NO 95-2797 950714; EP 680330 A1 EP 94-907261 940118, WO 94-US688 940118; JP 08506015 W JP 94-517167 940118, WO 94-US688 940118; CN 1096584 A CN 94-101086 940121; US 5629166 A Cont of US 93-8446 930122, US 94-288977 940811; AU 680236 B AU 94-60913 940118
FDT AU 9460913 A Based on WO 9416713; EP 680330 A1 Based on WO 9416713; JP 08506015 W Based on WO 9416713; AU 680236 B Previous Publ. AU 9460913, Based on WO 9416713
PRAI US 93-8446 930122; US 94-288977 940811
AB WO 9416713 A UPAB: 971113
Candidates for treatment with a therapeutic agent (A) specific for complexes (C) of HLA-C-clone 10 and a **MAGE-1** derived peptide are identified by (1) treating an abnormal cell sample from the patient with cytolytic T cells (CTC) specific for (C) and (2) detecting cell lysis to indicate that the subject is suitable for treatment.

Also new are (1) treatment of cellular abnormalities with CTC responsive to (C) or cell surfaces or with an agent inducing such CTC; (2) isolated CTC specific for (C) and (3) detecting abnormal cells having (C) on the surface by detecting lysis induced by these CTC.

USE - The cellular abnormality being treated is cancer, esp.

melanoma, but may also be non-proliferating cells with (C) on the surface, e.g. cells involved in autoimmune disorders. Detection of (C) can be used for diagnosis.

In an example, the melanoma cell line MZ2-MEL43 expresses a protein functionally equiv. (if not identical) to HLA-C-clone 10. A plasmid contg. cDNA for this protein and a second plasmic contg. cDNA for **MAGE-1**, -2 or -3 were used to transfect COS-7 cells. The transformants were then incubated with CTC of clone 81/12 (isolated from patient MZ2 and recognising MZ2-MEL.43) and after 24 hr. the concn. of tumour necrosis factor in the supernatant measured. A tumour rejection antigen derived from **MAGE-1** is presented by HLA-C-clone 10 and recognised by clone 81/12. Expression of MAGE-2 or -3 did not result in antigen presentation.

Dwg.0/1

L1 ANSWER 9 OF 9 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 94-035970 [05] WPIDS
DNC C94-016558
TI Monoclonal antibodies for diagnosis or therapy - directed against conjugate of MHC class I mol and peptide antigen.
DC B04 D16
IN HAEMMERLING, G
PA (DEKR-N) DEUT KREBSFORSCHUNGSZENTRUM
CYC 1
PI DE 4224542 A1 940127 (9405)* 3 pp
ADT DE 4224542 A1 DE 92-4224542 920724
PRAI DE 92-4224542 920724
AB DE 4224542 A UPAB: 940315

Producing monoclonal antibodies directed against a conjugate of a major histocompatibility complex class I mol. (I) and a peptide antigen (II) comprises (a) isolating (I); (b) inserting a (I)-encoding gene into the genome of a mouse to permit expression of the gene; (c) conjugating (I) with (II); (d) immunising the transformed mouse with the conjugate; (e) isolating spleen cells from the mouse; and (f) producing and opt. humanising monoclonal antibodies in known manner.

Process (I) may be isolated from CS7 AL/6 mouse RMA-S tumour cells or human EBV transformed cells, or may be isolated from tissue, or may be produced by recombinant DNA techniques. Step (b) may be omitted if (I) was isolated from the same mouse strain as that to be immunised. (II) may be a viral or tumour antigen, e.g, the human melanoma antigen **MAGE-1** or the tumour antigen produced by the HPV E6 or E7 oncogene.

USE - The antibodies are useful for diagnosis and therapy of tumours and infections, e.g, as a substitute for tumour-specific cytotoxic T cells.

Dwg.0/0

L8 ANSWER 1 OF 8 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 96-433764 [43] WPIDS
DNN N96-365449 DNC C96-136164
TI Anti-lung **tumour antigen** monoclonal antibody TB2A36C3 - produced by Epstein-Barr virus transformation of human lung cancer patient B-cells, useful in conjunction with other agents

for lysis of tumours.
DC B04 D16 S03
IN MUKERJEE, S
PA (MEDE-I) MEDENICA R D
CYC 69
PI WO 9628473 A1 960919 (9643)* EN 46 pp
RW: AT BE CH DE DK EA ES FI FR GB GR IE IT KE LS LU MC MW NL OA PT SD SE SZ UG
W: AL AM AT AU AZ BB BG BR BY CA CH CN CZ DE DK EE ES FI GB GE HU IS JP KE KG KP KR KZ LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK TJ TM TR TT UA UG UZ VN
AU 9653652 A 961002 (9703)
ADT WO 9628473 A1 WO 96-US3661 960318; AU 9653652 A AU 96-53652 960318
FDT AU 9653652 A Based on WO 9628473
PRAI US 95-405034 950316
AB WO 9628473 A UPAB: 961025
Novel monoclonal antibody (MAb) TB2A36C3 (pref. having the 134 residue amino acid light chain sequence given in the specification), has high specificity against lung **tumour antigens** and is produced by an Epstein-Barr virus (EBV) transformed TB945 human B-cell line. Also claimed are: (1) transformed human B-cell line immortalised by EBV (2) MAb produced by a cell as in (1); (3) human MAb which shows positive reactivity against small cell lung cancer (SCLC) and non-SCLC (NSCLC), and no reactivity against breast, ovary, **melanoma**, leiomyosarcoma and leukaemia/lymphoma cell lines; (4) MAb which specifically binds to a 32 **kD** mol. wt. antigen on NCIH 69 cell line, and a cluster of antigens from 28-106 **kD** in the NSCLC cell ine NCIH661; and (5) bioreagent for Ab assays comprising a pure peptide fragment F(ab)'2 of the MAb TB2A36C3.
USE - The MAb can be used for screening serum or tissue samples for a carcinoma-associated antigen (claimed). It can also be used, opt. with other agents, for the lysis of tumours in anti-tumour therapy, and for activating immune competent CD4 or CD8 cells in a patient' s blood system (claimed).
Dwg.0/10

L8 ANSWER 2 OF 8 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 94-025876 [03] WPIDS
CR 92-381780 [46]
DNC C94-011912
TI Inhibiting tumour growth and metastases with mullerian inhibiting substance - or its C-terminal fragment, opt. introduced by gene transfer, also for modulating, expression of class 1 MHC antigens, e.g. in AIDS patients.
DC B04 D16
IN BARKSDALE, E M; CHIN, T W; DONAHOE, P K; EPSTEIN, J; MACLAUGHLIN, D T; PARRY, R L; RAGIN, R C
PA (GEHO) GEN HOSPITAL CORP
CYC 20
PI WO 9400133 A1 940106 (9403)* EN 165 pp
RW: AT BE CH DE DK ES FR GB GR IE IT LU MC NL PT SE
W: CA JP
EP 646010 A1 950405 (9518) EN
R: CH DE DK ES FR GB IE IT LI NL PT SE
EP 646010 A4 970423 (9735)
US 5661126 A 970826 (9740) 63 pp

ADT WO 9400133 A1 WO 93-US5791 930618; EP 646010 A1 EP 93-916585 930618, WO 93-US5791 930618; EP 646010 A4 EP 93-916585 ; US 5661126 A CIP of US 89-299158 890119, CIP of US 91-683966 910412, CIP of US 92-901637 920619, Cont of US 93-7125 930121, US 94-271252 940707
FDT EP 646010 A1 Based on WO 9400133
PRAI US 93-7125 930121; US 92-901637 920619; US 89-299158 890119; US 91-683966 910412; US 94-271252 940707
AB WO 9400133 A UPAB: 971006
Tumour growth is inhibited by admin. of Muellerian inhibiting substance (MIS), or its C-terminal fragment, opt. used in conjunction with a chemotherapeutic agent.

Also new are (1) local inhibition of tumour growth by proteolytic activation (cleavage to 57 and 12.5 **kD** fragments) at the tumour site, (2) two DNA sequences (both 327 bp., reproduced in the specification) encoding bovine and human MIS C-terminal fragment (109 amino acids), etc. and (b) method for increasing (decreasing) expression of mRNA encoding a class I MHC protein in a cell by contact with MIS (or MIS antagonist).

USE - MIS inhibits both prim. tumour growth and metastasis, esp. of vulvar epidermoid and cervical carcinoma, endometrial or ovarian adenocarcinoma, ocular **melanoma** and tumours of the prostate, lymphoid tissue, breast, skin and germ cells. The use of MIS can allow the amt. of (I) needed for tumour treatment to be reduced. Increasing expression of MHC **antigens** suppresses metastases or **tumour** cells and improves immunedeficiency e.g. in AIDs patients, pref. when admin. together with interferon. Reducing MHC expression is esp. with epidermal growth factor and is used to reduce immunogenicity of transplants.
Dwg.0/27

L8 ANSWER 3 OF 8 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 94-010204 [02] WPIDS
DNN N94-008206 DNC C94-004140
TI Receptor protein for CSVTCG region of thrombospondin and antibodies or antisera against it - used as medicaments against, e.g. tumour cell adhesion and metastasis, thrombosis, and malaria, and as targetting agents.
DC B04 D16 S03
IN EYAL, J; HAMILTON, B K; TUSZYNSKI, G P; HAMILTON, B
PA (GRAC) GRACE & CO-CONN W R; (MEDI-N) MEDICAL COLLEGE PENNSYLVANIA; (MEDI-N) MEDICAL COLLEGE PA
CYC 20
PI EP 578342 A2 940112 (9402)* EN 16 pp
R: AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE
CA 2095404 A 931115 (9406)
US 5367059 A 941122 (9501) 12 pp
EP 578342 A3 940629 (9527)
JP 07138296 A 950530 (9530) 24 pp
ADT EP 578342 A2 EP 93-250135 930511; CA 2095404 A CA 93-2095404 930503; US 5367059 A US 92-883659 920514; EP 578342 A3 EP 93-250135 930511; JP 07138296 A JP 93-132353 930512
PRAI US 92-883659 920514
AB EP 578342 A UPAB: 940223
Isolated receptor protein has specific binding affinity for CSVTCG peptide of sequence Cys-Ser-Val-Thr-Cys-Gly.

Also claimed are (1) an antibody or antisera specifically reactive to the receptor protein; and (2) a carrier molecule and

biomedical device comprising the antibody or antisera.

USE/ADVANTAGE - CSVTCG receptor protein can be purified from cell extracts of tumour cell lines, pref. **melanoma** or lung cancer cell lines, by passing the extracts through a chromatographic column contg. immobilised CSVTCG peptides. The antibodies or antisera can be used to diagnose carcinoma by reacting them with a tissue sample. They can also be used in the prepn. of a medicament to inhibit tumour cell adhesion or metastasis, or for inhibiting thrombotic activity, atherosclerosis, malarial activity or angiogenesis. The carrier molecule may also comprise toxins, drugs, hormones or imaging agents and can be used to deliver these cpds. to tumour cells for diagnostic or therapeutic purposes. The receptor is specific for the CSVTCG region of thrombospondulin.

Dwg.0/2

L8 ANSWER 4 OF 8 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 92-433665 [52] WPIDS
DNN N92-330897 DNC C92-192572
TI New monoclonal antibody ME20 - specific for cell surface **protein** detectable on human **melanoma** **tumour** cells, for diagnosis and treatment of **melanoma.**
DC B04 D16 S03
IN HELLSTROM, I; HELLSTROM, K E; MARQUARDT, H
PA (BRIM) BRISTOL-MYERS SQUIBB CO
CYC 23
PI WO 9221767 A1 921210 (9252)* EN 38 pp
RW: AT BE CH DE DK ES FR GB GR IT LU MC NL SE
W: AU CA FI JP KR NO
AU 9221706 A 930108 (9315)
ZA 9204068 A 930224 (9316) 35 pp
PT 100568 A 930930 (9342)
ADT WO 9221767 A1 WO 92-US4451 920527; AU 9221706 A AU 92-21706 920527, WO 92-US4451 920527; ZA 9204068 A ZA 92-4068 920604; PT 100568 A PT 92-100568 920604
FDT AU 9221706 A Based on WO 9221767
PRAI US 91-710613 910605
AB WO 9221767 A UPAB: 940407

An antibody (Ab) or fragment that specifically reacts with a human **melanoma** cell surface protein (I) og mol.wt. 80-120 **kD** and which is present on human **melanoma** tumour cells but not normal cells, is new.

Also new are (1) a purified peptide contg. a region, corresp. to a domain of (I) or a fragment, which can specifically immunoreact with monoclonal Ab ME20 produced by the hybridoma ATCC HB10764 or HB10763; (2) an immunogenic compsn. comprising this peptide and a carrier or an immune responding agent; (3) a kit comprising the Ab or fragment and/or the peptide.

USE/ADVANTAGE - The Ab can be used to determine the presence and location of human **melanoma** cells. It may also be used to treat **melanoma** when administered with a chemotherapeutic agent. The peptide is used to induce an immune response directed towards human **melanoma**. The Ab can be used in in vivo or in vitro assay

Dwg.0/0

L8 ANSWER 5 OF 8 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD

AN 91-353726 [48] WPIDS
DNN N91-270898 DNC C91-152567
TI New 35 **KD tumour** associated **protein** **antigen** - used for immuno-diagnosis, immuno-prognosis and therapy of cancer.
DC B04 D16
IN GUPTA, R K; MORTON, D L; WONG, J H
PA (REGC) UNIV CALIFORNIA
CYC 33
PI WO 9117187 A 911114 (9148)*
RW: AT BE CH DE DK ES FR GB GR IT LU NL OA SE
W: AT BB BG BR CA FI HU JP KP KR LK MC MG MW NO PL RO SD SU
EP 529007 A1 930303 (9309) EN 23 pp
R: AT BE CH DE DK ES FR GB GR IT LI LU NL SE
JP 05506241 W 930916 (9342) 9 pp
ADT EP 529007 A1 EP 91-920948 910417, WO 91-US2638 910417; JP 05506241 W JP 91-508637 910417, WO 91-US2638 910417
FDT EP 529007 A1 Based on WO 9117187; JP 05506241 W Based on WO 9117187
PRAI US 90-510602 900418
AB WO 9117187 A UPAB: 930928
A purified **tumour**-associated antigenic **protein** has a molecular weight of 35 **kD** by SDS-PAGE under reducing conditions and specifically binds MAb JSI. Also claimed are reagents which specifically bind the protein, reagents which bind JSI, a hybridoma cell line producing JSI, nucleic acid encoding the protein and nucleic acid hybridising to the nucleic acid that encodes the protein.
USE/ADVANTAGE - The protein is used to detect subclinical cancer, to monitor a malignancy, as a vaccine for inducing an immune response to cancers e.g. **melanoma**, sarcoma and carcinoma, especially breast carcinoma. Cancer is detected by enhancing expression of the protein on cancer cells with IFN-gamma, contacting the protein with a reagent (e.g. an antibody) and detecting the presence of the reagent. The reagent may also be used to detect an immune complex containing the protein.
0/1

L8 ANSWER 6 OF 8 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 91-164362 [22] WPIDS
DNN N91-125906 DNC C91-071183
TI New urinary **tumour** associated **antigen** and its sub-units - used for immuno diagnosis, immuno prognosis and therapy of human cancer.
DC B04 D16 S03
IN EUHUS, D M; GUPTA, R K; MORTON, D L; GUPTA, R
PA (EUHU-I) EUHUS D M; (GUPT-I) GUPTA R K; (MORT-I) MORTON D L
CYC 34
PI WO 9106866 A 910516 (9122)* 86 pp
RW: AT BE CH DE DK ES FR GB GR IT LU NL OA SE
W: AT AU BB BG BR CA CH DE DK ES FI GB HU JP KP KR LK LU MC MG MW NL NO RO SD SE SU
AU 9168753 A 910531 (9135)
EP 498851 A1 920819 (9234) EN 86 pp
R: AT BE CH DE DK ES FR GB GR IT LI LU NL SE
JP 05505596 W 930819 (9338) 31 pp
WO 9106866 A3 910905 (9508)
AU 661816 B 950810 (9540)

EP 678744 A2 951025 (9547) EN 44 pp
R: AT BE CH DE DK ES FR GB GR IT LI LU NL SE
EP 498851 B1 960103 (9606) EN 42 pp
R: AT BE CH DE DK ES FR GB GR IT LI LU NL SE
DE 69024659 E 960215 (9612)
EP 678744 A3 951213 (9619)
ES 2084715 T3 960516 (9627)
US 5700649 A 971223 (9806) 35 pp

ADT EP 498851 A1 EP 90-917644 901031, WO 90-US6339 901031; JP 05505596 W WO 90-US6339 901031, JP 91-500470 901031; WO 9106866 A3 WO 90-US6339 901031; AU 661816 B AU 91-68753 901031; EP 678744 A2 EP 95-104918 901031; EP 498851 B1 EP 90-917644 901031, WO 90-US6339 901031; DE 69024659 E DE 90-624659 901031, EP 90-917644 901031, WO 90-US6339 901031; EP 678744 A3 EP 95-104918 901031; ES 2084715 T3 EP 90-917644 901031; US 5700649 A Div ex US 89-431533 891103, US 95-462264 950605

FDT EP 498851 A1 Based on WO 9106866; JP 05505596 W Based on WO 9106866; AU 661816 B Previous Publ. AU 9168753, Based on WO 9106866; EP 498851 B1 Based on WO 9106866; DE 69024659 E Based on EP 498851, Based on WO 9106866; EP 678744 A3 Related to EP 498851; ES 2084715 T3 Based on EP 498851

PRAI US 89-431533 891103; US 95-462264 950605

AB WO 9106866 A UPAB: 930928

A new pure antigenic polypeptide subunit of urinary **tumour** associated **antigen** (UTAA) is claimed. After reduction by beta-mercaptoethanol and separation by SDS-PAGE, the subunit has a molecular weight of 90-100 **kD**. Also claimed are: 1) A monoclonal antibody (MAb) produced by isolating antibody producing cells from an animal exposed to the polypeptide forming hybridomas between these cells and cancer cells, and selecting those producing the MAb. 2) Reagents reactive with antibodies that react with UTAA. 3) An epitope of UTAA on the 45 **kD** subunit. 4) An epitope of UTAA and the 120 **kD** subunit. 5) An epitope of UTAA on the 65 **kD** subunit. 6) Nucleic acid encoding the subunit; and 7) A nucleic acid probe which hybridises with the nucleic acid of (5). The MAb is FgG or IgM.

USE/ADVANTAGE - The MAb is used to detect subclinical cancer, UTAA on the tumour cells of a biopsy and breast or lung cancer. Low levels of UTAA may be detected. A vaccine against tumour cells is also claimed, comprising inactivated tumour cells with a UTAA and at least one **tumour**-associated **antigen** selected from GM-2, GD-2, foetal **antigen** or **melanoma** **tumour**-associated **antigen**. Cancers that can be vaccinated against include melanomas, sarcomas and carcinomas. @(86pp Dwg.No.1/17)@

L8 ANSWER 7 OF 8 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 89-270468 [37] WPIDS
DNC C89-119828
TI New continuous cell line for prodn. of monoclonal antibody - reacts with **melanoma**-associated **tumour** cell **antigen**, useful for diagnosis and differentiation of cancerous diseases.
DC B04 D16
IN KHAN, A
PA (WADL) WADLEY TECHNOLOGIES INC; (CYTO-N) CYTOCLONAL PHARM INC
CYC 1
PI US 4851510 A 890725 (8937)* 10 pp

US 5654408 A 970805 (9737) 11 pp
ADT US 4851510 A US 84-676839 841130; US 5654408 A Div ex US 84-676839 841130, US 89-341718 890421
FDT US 5654408 A Div ex US 4851510
PRAI US 84-676839 841130; US 89-341718 890421
AB US 4851510 A UPAB: 930923
A continuous cell line which produces antibodies, which specifically bind with at least one site on a **melanoma**-associated antigen, but are unreactive with normal human cells, breast or ovarian carcinoma cells or H-29 colon carcinoma cells as determined by indirect immunoflorescence, is claimed. The site on the antigen is not associated with normal human cell surface antigens and the **melanoma**-associated antigen is isolatable from G-361 **melanoma** cells, the isolated antigen having a mol. wt. of 105000 as determined by sodium dodecyl sulphate gel electrophoresis.

Also claimed is a continuous cell line producing antibodies specific against a **melanoma**-associated antigen of mol. wt. 38000. Cell line ATCC HB8672 is claimed, and the monoclonal antibody WI-MN-1 it produces.

Also claimed is an antibody produced by (a) obtg. a WI-MN-1 monoclonal antibody produced by cell line ATCC HB8672; (b) forming a complex comprising WI-MN-1 bound to a **tumour** cell **antigen** by contacting WI-MN-1 with a **tumour** cell; (c) isolating the **tumour** cell **antigen** which combines with the WI-MN-1 monoclonal antibody from the complex; (d) injecting the **tumour** cell **antigen** into an animal; (e) selecting and isolating an antibody-forming cell from the animal, the cell capable of producing antibodies specifically binding the **tumour** cell **antigen**; (f) fusing the antibody-forming cell with a myeloma cell to form a hybridoma; and (g) causing the hybridoma to make the claimed antibody.

Alternatively, the isolated **tumour antigen** is injected into an animal, circulating fluids are collected and the antibody directed against the antigen is isolated.

USE/ADVANTAGE - The WI-MN-1 antibody may be used for diagnosis and differentiation of cancerous diseases.

0/0

L8 ANSWER 8 OF 8 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 86-163227 [26] WPIDS
DNC C86-069813
TI Antibody for differentiating between malignant and benign tumours - obtd. by contacting malignant tumour cells with gamma-globulin to induce antigenicity then using to immunise animal etc..
DC B04 D16
IN FERRONE, S; MATSUI, M
PA (SUNR) SUNTORY LTD
CYC 13
PI EP 185135 A 860625 (8626)* EN 15 pp
R: AT BE CH DE FR GB IT LI LU NL SE
JP 61148199 A 860705 (8633)
US 5126262 A 920630 (9229) 7 pp
ADT EP 185135 A EP 85-110206 850814; JP 61148199 A JP 85-217741 850930; US 5126262 A Cont of US 84-684262 841220, Cont of US 88-251980 880928, US 91-657253 910215
PRAI US 84-684262 841220
AB EP 185135 A UPAB: 930922

Antibody specifically differentiating between a malignant and a benign tumour is new when obtd. by: (a) contacting cultured malignant tumour cells with a cytokine or lymphokine antitumour agent (I) to enhance or induce the antigenicity of the cells specific to the malignant tumour; (b) immersing an animal with these cells; (c) obtg. spleen cells from the animal; (d) fusing the spleen cells with immortal cells for hybridoma prodn.; (e) cloning the fused cells for selection of hybridoma cells producing an antibody specific to the malignant tumour; and (f) collecting the antibody from a culture of the selected hybridoma cells.

USE/ADVANTAGE - The antibody is useful for the diagnosis and treatment of human tumour cells and tissues, and esp. for distinguishing malignant melanomas from benign **tumours**. The target **antigen** does not have to be purified and large amts. of homogeneous antibody can be produced.

0/0

```
=> fil biosis

FILE 'BIOSIS' ENTERED AT 09:28:46 ON 07 APR 1998
COPYRIGHT (C) 1998 BIOSIS(R)

FILE COVERS 1969 TO DATE.
CAS REGISTRY NUMBERS AND CHEMICAL NAMES (CNs) PRESENT
FROM JANUARY 1969 TO DATE.

RECORDS LAST ADDED: 6 April 1998 (980406/ED)
CAS REGISTRY NUMBERS (R) LAST ADDED: 6 April 1998 (980406/UP)


=> d his


     (FILE 'BIOSIS' ENTERED AT 09:23:41 ON 07 APR 1998)
                DEL HIS Y
L1           119 S MAGE 1
L2        328823 S SEQUENCE?
L3            19 S L1 AND L2
L4            13 S L1 AND MONOCLONAL
L5            17 S L3 NOT L4
L6             2 S 11540 OR HB11540
L7         26284 S (TUMOR# OR TUMOUR#) (4A) (PROTEIN# OR ANTIGEN#)
L8          1541 S L7 AND MELANOMA?
L9          3105 S (46 OR 34 ) (2W) (KD OR KDA OR KDS OR KILODALTON# OR KI
L10            7 S L9 AND L8
L11            4 S L1 AND L9
L12            8 S L10 OR L11

     FILE 'BIOSIS' ENTERED AT 09:28:46 ON 07 APR 1998

=> d bib ab l4 1-13;d bib ab l5 1-17;d bib ab st 1-8 l12


L4   ANSWER 1 OF 13  BIOSIS  COPYRIGHT 1998 BIOSIS
AN  97:503587  BIOSIS
DN  99802790
TI  Immunohistochemical detection of MAGE tumor-associated antigens in
```

esophageal squamous cell carcinoma.
AU Schaefer C; Noppen S; Nuernberger H-R; Gudat F; Kocher T; Luescher U; Zuber M; Spagnoli G C; Heberer M
CS Dep. Surg. Res., ZLF, Hebelstrasse 20, CH-4031 Basel, Switzerland
SO Oncology Reports 4 (6). 1997. 1289-1293. ISSN: 1021-335X
LA English
AB Genes of the MAGE family encode tumor-specific antigens recognized by cytotoxic T-lymphocytes in a variety of neoplasms. We investigated the protein expression of these antigens as related to the gene expression, in esophageal squamous cell carcinoma by using **monoclonal** antibodies recognizing MAGE gene products. Esophageal squamous cell carcinomas were found to express both **MAGE-1** (4 out of 15 samples) and MAGE-3 (7 out of 15 samples) genes, by RT-PCR. Immunoblotting revealed **MAGE-1** and MAGE-3 gene products in 2 and 6 out of 15 samples, respectively. Immunohistochemistry performed on 12 samples showed **MAGE-1** protein expression, limited to single tumor cells, in 2 cases. MAGE-3 gene product was detectable in 7 cases: in 5 of them over 50% of neoplastic cells were positive. Considering the high percentages of tumor cells expressing MAGE-3 antigen, the use of epitope-based vaccines could be envisaged in patients displaying appropriate HLA-class I phenotypes.

L4 ANSWER 2 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 97:503259 BIOSIS
DN 99802462
TI Heterogeneity of melanoma antigen-1 (**MAGE-1**) gene and protein expression in malignant melanoma.
AU Zuber M; Spagnoli G C; Kocher T; Lusher U; Schaefer C; Noppen C; Gudat F; Harder F; Heberer M
CS Dep. Surg., Univ. Basel, Spitalstr. 21, CH-4031 Basel, Switzerland
SO European Surgical Research 29 (5). 1997. 403-410. ISSN: 0014-312X
LA English
AB Objective: The authors' objective is to identify **MAGE-1** tumor antigen in clinical melanoma specimens and to verify the extent of its expression in tumors where evidence of specific gene transcripts can be obtained. Background data: The **MAGE-1** gene encodes a tumor-associated antigen that can be recognized by specific cytotoxic T lymphocytes. Transcription of the **MAGE-1** gene has previously been demonstrated in various malignancies, but the production of the specific gene product and its distribution in neoplastic tissues have not yet been addressed. Methods: Total cellular mRNA was extracted from six melanoma biopsies, reverse-transcribed and tested in 25-45 cycles of reverse polymerase chain reaction (rtPCR) in the presence of primers' pairs specific for the beta-actin-positive control gene and for the **MAGE-1**-encoding gene. Concurrently, portions of these specimens were lysed and probed for **MAGE-1** protein by immunoblotting. Additional material from the same biopsies was analyzed following immunohistological staining with **MAGE-1**-specific **monoclonal** antibodies. Results: **MAGE-1** gene transcription could be demonstrated following 25 cycles of rtPCR in one out of six biopsies and in three more following 35 cycles of rtPCR. 2/6 samples were negative even after 45 cycles of rtPCR. **MAGE-1** protein production could be detected by immunoblotting in the lysates from biopsies showing evidence of specific gene transcription. Cells

positive for **MAGE-1** protein expression could be identified by immunohistochemistry on snap-frozen sections in three of the four tumors displaying specific transcripts. Distribution of positivity ranged between focal cellular areas and single positive cells in the different tumors. Conclusions: The **MAGE-1** tumor antigen can be detected by specific **monoclonal** antibodies in clinical tumor specimens. The pattern of positivity observed in samples showing evidence of **MAGE-1** gene expression suggests a relevant heterogeneity regarding **MAGE-1** antigen production within individual tumor specimens.

L4 ANSWER 3 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:563309 BIOSIS
DN 99292665
TI The tumour-associated antigen **MAGE-1** is detectable in formalin-fixed paraffin sections of malignant melanoma.
AU Gudat F; Zuber M; Durmuller U; Kocher T; Schaefer C; Noppen C; Spagnoli G
CS Inst. Pathol., Univ. Basel, Schonbeinstrasse 40, CH-4003 Basel, Switzerland
SO Virchows Archiv 429 (2-3). 1996. 77-81. ISSN: 0945-6317
LA English
AB The **MAGE-1** gene encodes a protein encompassing a HLA-A1-restricted target epitope for cytolytic T lymphocytes. **Monoclonal** antibodies directed against the **MAGE-1** protein were tested for usage in immunohistology of routine pathology material. Seven formalin-fixed, paraffin-embedded malignant melanomas were studied by the Avidin-Biotin complex (ABC) method with or without different antigen retrieval methods. Native, frozen tissues from the same tumours were used to validate the results by immunohistochemistry on frozen sections, by PCR for mRNA and by protein demonstration in tissue extracts using western blotting. Of 4 **monoclonal** antibodies tested, mAB 34B and mAB 77B were highly efficient in detecting **MAGE-1** protein in deparaffinized sections with the regular ABC method after microwave pretreatment. In a series of an additional 28 patients 75% expressed **MAGE-1**, 50% in a substantial proportion. Follow-up studies in 6 patients indicate that the expression pattern remains stable but may change substantially within a short range. Immunohistology is thus a rapid and well-established method that might be used to select and monitor HLA-A1 positive patients with malignant melanoma and other candidate tumours for **MAGE-1**-directed immuno-therapy.

L4 ANSWER 4 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:415356 BIOSIS
DN 99137712
TI **Monoclonal** antibodies against recombinant-**MAGE-1** protein identify a cross-reacting 72-kDa antigen which is co-expressed with **MAGE-1** protein in melanoma cells.
AU Carrel S; Schreyer M; Spagnoli G; Cerottini J-C; Rimoldi D
CS Ludwig Inst. Cancer Res., Ch. des Boveresses 155, 1066 Epalinges, Switzerland
SO International Journal of Cancer 67 (3). 1996. 417-422. ISSN: 0020-7136

LA English
AB The MAGE-I gene codes for tumor-associated peptides recognized by cytolytic T lymphocytes in association with MHC-class-I molecules such as HLA-A I and HLA-Cw16. In the course of a study aiming at the immunohistochemical detection of the MAGE-I gene product in tumor samples, 2 mouse **monoclonal** antibodies (MAbs) directed against a full-length recombinant MAGE-I fusion protein were found to react strongly not only with the 46-kDa MAGE-I protein, but also with a 72-kDa product in immunoblots of lysates obtained from several MAGE1-mRNA-positive melanoma cell lines. Pre-incubation of the antibodies with the recombinant MAGE-I fusion protein abolished their reactivity both with MAGE-I protein and with the 72-kDa product, thus confirming the occurrence of antigenic determinant(s) shared by the 2 proteins. The 72-kDa protein is not an alternative product of MAGE-I, since it was still detected in lysates of a MAGE-I loss variant derived from a MAGE-I-positive melanoma cell line. Moreover, the 72-kDa protein does not appear to be a product of the other members of the MAGE gene family known to be expressed in tumors (such as MAGE-2, -3, -4 and -12). Interestingly, expression of the 72-kDa protein was found to be correlated with that of MAGE-I protein. Thus, in 30 tumor cell lines analyzed by immunoblotting and RT-PCR, the 72-kDa protein was never detected in MAGE-I-mRNA-negative cell lines, while it was co-expressed with MAGE-I protein in 12 out of 15 cell lines expressing MAGE-I. Furthermore, the 72-kDa protein was detected in lysates of human testis, the only normal tissue known to express MAGE-I. Finally, treatment of **MAGE- 1** -mRNA-negative cell lines with 5-Aza-2'-deoxycytidine, a hypomethylating agent known to induce MAGE- I expression, resulted in the expression of the 72-kDa protein. Taken collectively, these findings suggest that expression of the gene encoding the 72-kDa protein identified in this study through antigenic determinant(s) shared with MAGE-I protein is regulated in a way similar to that of MAGE-I.

L4 ANSWER 5 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:257272 BIOSIS
DN 98813401
TI **Monoclonal** antibodies against full-length **MAGE- 1** protein identify a crossreacting 72 kDa antigen which is coexpressed with **MAGE-1** melanoma cells.
AU Carrel S; Salvi S; Hartmann F; Rimoldi D; Schreyer M; Spagnoli G
CS Ludwig Inst. Cancer Res., Lausanne Branch, Univ. Lausanne, 1066 Epalinges, Switzerland
SO 87th Annual Meeting of the American Association for Cancer Research, Washington, D.C., USA, April 20-24, 1996. Proceedings of the American Association for Cancer Research Annual Meeting 37 (0). 1996. 465. ISSN: 0197-016X
DT Conference
LA English

L4 ANSWER 6 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:46201 BIOSIS
DN 98618336
TI A 72kDa protein coexpressed in melanoma cells with **MAGE- 1**, revealed by anti-recombinant **MAGE-1 monoclonal** antibodies.
AU Carrel S; Hartmann F; Salvi S; Spagnoli G; Schreyer M; Rimoldi D

CS Ludwig Inst. Cancer Research, Lausanne Branch, Univ. Lausanne, 1066 Epalinges, Switzerland
SO Fifth International Conference of Anticancer Research, Corfu, Greece, October 17-22, 1995. Anticancer Research 15 (5A). 1995. 1672. ISSN: 0250-7005
DT Conference
LA English

L4 ANSWER 7 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:551277 BIOSIS
DN 98565577
TI Establishment of an enzyme-linked immunosorbent assay (ELISA) for measuring cellular MAGE-4 protein on human cancers.
AU Shichijo S; Tsunosue R; Kubo K; Kuramoto T; Tanaka Y; Hayashi A; Itoh K
CS Dep. Immunol., Kurume Univ. Sch. Med., Kurume 830, Japan
SO Journal of Immunological Methods 186 (1). 1995. 137-149. ISSN: 0022-1759
LA English
AB The MAGE genes encoding tumor-rejection antigens are expressed on various human cancers. An enzyme-linked immunosorbent assay (ELISA) was established for measuring cellular MAGE-4 protein (MAGE-4a and/or -4b) expressed on human tumor cells using a **monoclonal** antibody (mAb) and polyclonal Ab to recombinant MAGE-4b protein. Both the R5 mAb (IgG1) and the polyclonal Ab recognized a 45 kDa protein in extracts of MAGE-4 mRNA positive cancers, and showed no apparent cross-reactivity to the other MAGE gene products (**MAGE-1**, -2, -3, -6, and -12) by the immunoblot analyses. The R5 mAb and the polyclonal Ab primarily recognized one (the position 119-133) and two oligopeptides (the positions 119-133 and 259-273), respectively, among a series of 31 different MAGE-4b oligopeptides. The amino acid sequences of these two peptides were identical to those of MAGE-4a and -4b, but differed from those of all the other MAGE proteins (**MAGE-1**, -2, -3, -6, and -12). Substitution of glycine for amino acid in position 123 (arginine, R), 124 (lysine, K), 126 (R) or 128 (K) in a MAGE-4b oligopeptide of the position 119-132 severely decreased the reactivity of the R5 mAb to the oligopeptide. This ELISA also showed no apparent cross-reactivity with the other MAGE gene products (**MAGE-1**, -2 -3, -6, and -12). The minimum detectable level of MAGE-4 protein was determined to be 10 pg/well (100 pg/ml). The results suggest that this ELISA is a reliable and quantitative method to measure cellular MAGE-4 protein that is a potential target molecule for specific immunotherapy of human cancers.

L4 ANSWER 8 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:487231 BIOSIS
DN 98501531
TI Expression of the MAGE gene family in human lymphocytic leukemia.
AU Schichijo S; Tsunosue R; Masuoka K; Natori H; Tamai M; Miyajima J; Sagawa K; Itoh K
CS Dep. Immunol., Kurume Univ. Sch. Med., 67-Asahi-machi, Kurume, Fukuoka 830, Japan
SO Cancer Immunology Immunotherapy 41 (2). 1995. 95-103. ISSN: 0340-7004
LA English
AB The MAGE gene family, encoding tumor-rejection antigens recognized by

cytotoxic T lymphocytes, is frequently expressed in human solid cancers. However, its expression in leukemia has not been well studied. We have investigated MAGE gene expression at the mRNA level in human leukemia. The MAGE gene family was expressed in 17 of 34 (50%) examples of T cell leukemia (12/21 patients' peripheral blood mononuclear cells and 5/13 cell lines), in 7 of 16 (44%) cases of B cell leukemia (1/8 and 6/8 respectively), but in none of 23 myelomonocytic leukemia cases (0/16 and 0/7), as evaluated by the primers common to the **MAGE-1**, -3, -4 (-4a and/or -4b), and -6 genes and the semi-quantificative reverse transcription/polymerase chain reaction method. None of a panel of normal lymphoid cells expressed the MAGE gene family. As revealed by the primers specific for each of the MAGE genes, the **MAGE-1**, -2, -3, -4 or -6 gene was expressed in 8, 8, 6, 2, or 6 respectively out of 23 types of leukemia cell lines. Expression of the **MAGE-1** protein in both the cell lines and patients' cells was confirmed by immunoblot analysis with the polyclonal antibody to recombinant **MAGE-1** protein. Cellular MAGE-4 protein in the cell lines was measured by an enzyme-linked immunosorbent assay with the polyclonal and **monoclonal** antibodies to recombinant MAGE-4b protein. In summary, the MAGE gene family was found to be expressed in the substantial proportion of T cell leukemias, but in no case of myelomonocytic leukemia. Antigens coded by the MAGE gene family could be important molecules for understanding specific immunity against lymphocytic leukemia.

L4 ANSWER 9 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:411919 BIOSIS
DN 98426219
TI Detection of MAGE-4 protein in lung cancers.
AU Shichijo S; Hayashi A; Takamori S; Tsunosue R; Hoshino T; Sakata M; Kuramoto T; Oizumi K; Itoh K
CS Dep. Immunol., Kurume Univ. Sch. Med., 67 Asahi-machi, Kurume 830, Japan
SO International Journal of Cancer 64 (3). 1995. 158-165. ISSN: 0020-7136
LA English
AB Expression of genes of the MAGE family, which encode tumor-rejection antigens recognized on HLA-A1 and -Cw1601 by cytotoxic T lymphocytes (CTL), was investigated in lung cancers at the mRNA (**MAGE-1**,-2,-3/-6, and -4 (4a and/or 4b)) and protein (MAGE-4) levels. **MAGE-1**, -2, -3/-6 and -4 genes were expressed, respectively, at the mRNA level in 6, 7, 20 and 7 of 53 lung cancers (SO non-small-cell lung cancers and 3 small-cell lung cancers) by the reverse transcription-polymerase chain reaction (RT-PCR) method. Polyclonal antibody (Ab) and **monoclonal** antibody (MAb) against recombinant MAGE-4b protein were developed to detect MAGE-4 protein. Both the polyclonal Ab and the RS MAb recognized a 45-kDa protein in extracts of MAGE-4 mRNA-positive lung cancers, and showed no apparent cross-reactivity with the other MAGE gene products except with MAGE-4a by immunoblot analyses and transfection experiments. MAGE-4 protein was detected on 13 of 44 (30%) lung cancers (18 to 55,989 pg/mg) by ELISA with the polyclonal Ab and RS MAb. These 13 lung cancers consisted of 6 of 6 MAGE-4 mRNA-detectable and 7 of 38 MAGE-4 mRNA undetectable lung cancers. Histologically, these comprised 7 of 10 squamous-cell carcinomas, 4

of 30 adenocarcinomas and 2 of 3 small-cell lung cancers. The proportions of MAGE gene-positive samples, at both the mRNA and protein levels, correlated with the size of the primary tumors and with regional node involvement. These results should provide important information on specific immunotherapy of lung cancers using MAGE gene products.

L4 ANSWER 10 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:317017 BIOSIS
DN 98331317
TI Identification and intracellular location of MAGE-3 gene product.
AU Kocher T; Schultz-Thater E; Gudat F; Schaefer C; Casorati G; Juretic A; Willimann T; Harder F; Heberer M; Spagnoli G C
CS Surg. Res. Lab., 20 Hebelstrasse 4031 Basel, Switzerland
SO Cancer Research 55 (11). 1995. 2236-2239. ISSN: 0008-5472
LA English
AB The human MAGE-3 gene encodes a melanoma antigenic epitope recognized by specific cytotoxic T lymphocytes, but its gene product has not been identified thus far. We produced a recombinant MAGE-3 gene product by expression cloning of the entire reading frame in the context of a fusion protein characterized by a 10-histidine tail, allowing purification by metal chelation on a nickel Sepharose column. The semipurified product was used to generate MAGE-3-specific **monoclonal** antibodies. One reagent could identify by immunoblotting the native MAGE-3 gene product as a M, 48,000 protein in lysates of cell lines showing evidence of MAGE-3 gene expression. No apparent cross-reactivity with recombinant or native **MAGE-1** gene product was observed. Immunohistochemistry shows that, closely resembling the **MAGE-1** gene product, MAGE-3 is a cytoplasmic protein.

L4 ANSWER 11 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:187036 BIOSIS
DN 98201336
TI Immunohistochemical localization of **MAGE-1** in melanoma cell lines by **monoclonal** anti-**MAGE-1** antibodies.
AU Carrel S; Hartmann F; Salvi S; Spagnoli G; Schreyer M; Rimoldi D
CS Ludwig Inst. Cancer Res., Lausanne Branch, Univ. Lausanne, 1066 Epalinges, Switzerland
SO Eighty-sixth Annual Meeting of the American Association for Cancer Research, Toronto, Ontario, Canada, March 18-22, 1995. Proceedings of the American Association for Cancer Research Annual Meeting 36 (0). 1995. 479. ISSN: 0197-016X
DT Conference
LA English

L4 ANSWER 12 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 94:548546 BIOSIS
DN 98008094
TI **MAGE-1** gene product is a cytoplasmic protein.
AU Schultz-Thater E; Juretic A; Dellabona P; Luscher U; Siegrist W; Harder F; Heberer M; Zuber M; Spagnoli G C
CS Z.L.F., Surgical Res. Lab., 20 Hebelstrasse, CH-4031 Basel, Switzerland
SO International Journal of Cancer 59 (3). 1994. 435-439. ISSN: 0020-7136

LA English
AB **MAGE-1** gene encodes a human melanoma antigen, recognized by syngeneic cytotoxic T lymphocytes (CTL). **MAGE -1** transcripts are also detectable in breast cancers, in non-small-cell lung carcinomas and in central nervous system tumors. In order to identify, in cellular preparations, the protein encompassing the antigenic peptide, we generated a panel of **monoclonal** antibodies (MAbs) against the **MAGE-1** gene product by using, as immunogen, a full-length recombinant preparation (rMAGE1), obtained through expression cloning of the relevant gene in E. coli. Four reagents were obtained recognizing both rMAGE-1 and the 46-kDa native protein in cell lines expressing **MAGE-1** mRNA. No positivity could be detected in **MAGE-1**-mRNA-negative melanoma lines. No surface labelling of **MAGE-1**-positive cell lines could be observed. In contrast, on permeabilization of MZ2 melanoma cells, all 4 MAbs induced efficient staining, as detected by cytofluorography. Fluorescence microscopy shows that **MAGE-1** gene product is a cytoplasmic protein clustered in paranuclear organelle-like structures. Thus, **MAGE-1** protein location closely resembles that of P91A and P198 murine-tumor antigens.

L4 ANSWER 13 OF 13 BIOSIS COPYRIGHT 1998 BIOSIS
AN 94:129901 BIOSIS
DN 97142901
TI Identification of the **MAGE-1** gene product by **monoclonal** and polyclonal antibodies.
AU Chen Y-T; Stockert E; Chen Y; Garin-Chesa P; Rettig W J; Van Der Bruggen P; Boon T; Old L J
CS Ludwig Inst. Cancer Res., New York Unit, New York Hosp.-Cornell Med. Cent., New York, NY 10021, USA
SO Proceedings of the National Academy of Sciences of the United States of America 91 (3). 1994. 1004-1008. ISSN: 0027-8424
LA English
AB The human **MAGE-1** gene encodes a melanoma peptide antigen recognized by autologous cytotoxic T lymphocytes. To produce antibodies against the **MAGE-1** gene product, several approaches were taken. Three oligopeptides were synthesized based on predicted **MAGE-1** amino acid sequences and were used to generate rabbit anti-peptide antisera. In addition, a truncated **MAGE-1** cDNA was cloned into an Escherichia coli expression vector, and recombinant protein was produced and purified. All three rabbit anti-peptide antisera showed reactivity against the immunizing peptide, and one reacted with the recombinant **MAGE-1** protein by immunoblotting, but none reacted with cell lysates from **MAGE-1** mRNA-positive cells. The recombinant **MAGE-1** protein was then used for the generation of mouse **monoclonal** and rabbit polyclonal antibodies. One IgG1 **monoclonal** antibody, MA454, as well as rabbit polyclonal antisera recognized a 46-kDa protein in extracts of **MAGE-1** mRNA-positive melanoma cell lines. The antibodies showed no apparent crossreactivity with products of the closely related MAGE-2 and MAGE-3 genes. Serological typing of normal and tumor cell lysates was in full agreement with mRNA analysis, showing expression of **MAGE-1** protein in **MAGE-1**

mRNA-positive testis and a subset of melanomas but not in **MAGE-1** mRNA-negative normal or tumor tissues. Transfection of the **MAGE-1** gene into a **MAGE-1** mRNA-negative melanoma cell line resulted in the expression of the 46-kDa protein, confirming the identity of this protein as the **MAGE-1** gene product.

L5 ANSWER 1 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 98:93450 BIOSIS
DN 01093450
TI Two members of the human MAGEB gene family located in Xp21.3 are expressed in tumors of various histological origins.
AU Lurquin C; De Smet C; Brasseur F; Muscatelli F; Martelange V; De Plaen E; Brasseur R; Monaco A P; Boon T
CS Ludwig Inst. Cancer Res., 74 avenue Hippocrate, UCL 74.59, B-1200 Brussels, Belgium
SO Genomics 46 (3). 1997. 397-408. ISSN: 0888-7543
LA English
AB Genes of the MAGE family direct the expression of tumor antigens recognized on a human melanoma by autologous cytolytic T lymphocytes. Twelve closely related MAGE genes are located in the Xq28 region. These genes share 60-98% nucleotide identity in their coding region. The presence of homologous genes in a region of Xp21.3 has been reported previously. We obtained the complete **sequence** of a 42-kb stretch of this region. It contains four MAGE-related genes, which we propose to name MAGE-B1, B2, B3, and B4 (HGMW-approved symbols MAGEB1, MAGEB2, MAGEB3, and MAGEB4). The coding regions of these genes share 66-81% nucleotide identity and show 4563% identity with those of the MAGE genes located in Xq28. Like the MAGE genes located in Xq28, the RAGEB genes are silent in normal tissues with the exception of testis. Like **MAGE-1**, 2, 3, 4, 6 and 12 (HGMW-approved symbols MAGEA1, 2, 3, 4, 6, and 12), genes MAGE-B1 and MAGE-B2 are expressed in a significant fraction of tumors of various histological types. The transcription of MAGE-B1 and MAGE-B2 can be induced by 5-aza-2'-deoxycytidine, suggesting that the activation of these genes in tumors results from a demethylation process.

L5 ANSWER 2 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 97:180328 BIOSIS
DN 99472041
TI Alternative promoters of gene MAGE4a.
AU De Plaen E; Naerhuyzen B; De Smet C; Szikora J-P; Boon T
CS Ludwig Inst. Cancer Res., Brussels Branch, 74 avenue Hippocrate, UCL 74.59, B1200 Brussels, Belgium
SO Genomics 40 (2). 1997. 305-313. ISSN: 0888-7543
LA English
AB Gene MAGE-4 (HGMW-approved symbol MAGE4) is expressed in several types of tumors, but not in normal tissues, except testis and placenta. The 5' end of this gene contains eight homologous exons spread over a 5.8-kb region. These exons are alternatively spliced to a unique second exon and a unique third exon, which encodes a protein of 317 amino acids. The analysis of transcripts found in testis, placenta, and a sarcoma cell line showed that each of the alternative first exons is used in at least one of these tissues. Various regions

of the promoter of the fifth alternative exon (1.5) were cloned in a luciferase reporter plasmid, and the constructs were transfected in a sarcoma cell line that expresses MAGE-4. Two Ets motifs located between positions -70 and -29 relative to the transcription start site were found to drive 55% of the promoter activity. A region containing a Sp1 consensus binding site located upstream of the two Ets motifs was found to be responsible for 44% of the transcriptional activity. MAGE-4a promoters 1.4 and 1.6, which also contain the Sp1 and the two Ets binding motifs, supported a level of transcription comparable to that of promoter 1.5, whereas promoter 1.1, which contains only one Ets binding site, was sixfold less active. In line with observations made with gene **MAGE-1** (HGMW-approved symbol MAGE1), we found that promoter 1.5 stimulated a high level of transcription in a melanoma cell line that does not express MAGE-4. This suggests that the tumor-specific expression of MAGE genes is not determined by the presence of specific transcription factors.

L5 ANSWER 3 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 97:19666 BIOSIS
DN 99318869
TI Methylated CpG points identified within **MAGE-1** promoter are involved in gene repression.
AU Serrano A; Gracia A; Abril E; Garrido F; Ruiz-Cabello F
CS Serv. Analisis Clinicos Inmunologia, Hospital Universitario Virgen de las Nieves, Univ. Granada, Avda. Constitucion s/n, 18014 Granada, Spain
SO International Journal of Cancer 68 (4). 1996. 464-470. ISSN: 0020-7136
LA English
AB The **MAGE-1** gene, expressed in some tumors of different histological origins, codes for a tumor antigen recognized by cytotoxic T lymphocytes. The gene is not expressed in normal tissues with the exception of testes. The present study was designed to investigate the relationship between methylation of the **MAGE-1** promoter and inactivation of the **MAGE-1** gene. We examined the extent to which **MAGE-1** B'B promoter **sequences** are methylated in tumor-cell lines, in order to determine whether methylation correlates with **MAGE-1** expression. Using methylation-sensitive restriction analysis followed by polymerase chain reaction (PCR), we found an inverse correlation between methylation of the **MAGE-1** B'B region and **MAGE-1** expression. An unmethylated state was identified in DNA from sperm and some tumor-cell lines of different origins. In contrast, a hypermethylation state was found in leukocytes and other **MAGE-1** non-expressing cells. Furthermore, treatment with 5-aza-2'-deoxycytidine, a demethylating agent, induced **MAGE-1** expression in tumor-cell lines in which we found no direct relation between transcriptional activity of the B'B region and **MAGE-1** expression. Binding of the nuclear factors to the B'-methylated probe was strongly inhibited, indicating that methylation of cytosine interferes directly in the binding of transcriptional factors.

L5 ANSWER 4 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:511599 BIOSIS

DN 99233955
TI Conserved TCR usage by HLA-Cw* 1601-restricted T cell clones recognizing melanoma antigens.
AU Farina C; Van Der Bruggen P; Boel P; Parmiami G; Sensi M
CS Div. Exp. Oncol. D, Istituto Nazionale per lo Studio e la Cura dei Tumori, Via Venezian 1, 20133 Milano, Italy
SO International Immunology 8 (9). 1996. 1463-1466. ISSN: 0953-8178
LA English
AB In this study we determined TCR alpha and beta chain nucleotide **sequences** of HLA-Cw*1601-restricted cytotoxic T lymphocyte (CTL) clones obtained from the peripheral blood lymphocytes (PBL) of a melanoma patient. These clones were previously shown to be involved in the recognition of melanoma-associated antigenic epitopes SAYGEPRKL and AARAVFLAL encoded by gene **MAGE-1** and BAGE respectively. All (3/3) anti-**MAGE-1** CTL clones displayed TCRBV5 usage and one clonotype was found twice, gt 1 year apart, in patient's PBL. Two out of three anti-BAGE CTL clones showed the same TCRAV/AJ and TCRBV/BJ combinations and differed in the a chain CDR3 for two residues and in the beta chain CDR3 for a single nucleotide which, however, did not change translation. These results suggest a pattern of TCR conservation in CTL selected for recognition of **MAGE-1** or BAGE peptides on the autologous melanoma.

L5 ANSWER 5 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:383076 BIOSIS
DN 99105432
TI Identification of peptide epitopes of **MAGE-1**, -2, -3 that demonstrate HLA-A3-specific binding.
AU McIntyre C A; Rees R C; Platts K E; Cooke C J; Smith M O; Mulcahy K A; Murray A K
CS Inst. Cancer Stud., Univ. Sheffield Med. Sch., Beech Hill Road, Sheffield S10 2RX, UK
SO Cancer Immunology Immunotherapy 42 (4). 1996. 246-250. ISSN: 0340-7004
LA English
AB The MAGE gene family of tumour antigens are expressed in a wide variety of human cancers. We have identified 43 nonamer peptide **sequences**, from **MAGE-1**, -2 and -3 proteins that contain binding motifs for HLA-A3 MHC class I molecules. The T2 cell line, transfected with the cDNA for the HLA-A3 gene, was used in a MHC class I stabilisation assay performed at 37 degree C and 26 degree C. At 37 degree C, 2 peptides were identified that stabilised HLA-A3 with high affinity (fluorescence ratio, FR gt 1.5), 4 peptides with low affinity (FR 1.11-1.49) and 31 peptides that did not stabilise this HLA haplotype (FR lt 1.1). At 26 degree C, 12 peptides were identified that stabilised HLA-A3 with high affinity, 8 peptides with low affinity and 17 peptides that did not stabilise this HLA haplotype. Two peptides stabilised HLA-A3 at both temperatures. Small changes in one to three amino acids at positions distinct from the anchor residues altered peptide affinity. Data were compared to a similar study in which a peptide competition assay was used to investigate **MAGE-1** peptide binding to several HLA haplotypes. This study demonstrates that anchor residues do not accurately predict peptide binding to specific HLA haplotypes, changes in one to three amino acids at positions distinct from anchor residues influence peptide binding and alternative methods of

determining peptide binding yield different results. We are currently investigating the ability of these peptides to induce antitumour cytotoxic T lymphocyte activity as they may be of potential therapeutic value.

L5 ANSWER 6 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:333217 BIOSIS
DN 99055573
TI Epitope specificity of anti-HIV antibodies in human and murine autoimmune diseases.
AU Fraziano M; Montesano C; Lombardi V R M; Sammarco I; De Pisa D; Mattei M; Valesini G; Pittoni V; Colizzi V
CS Dep. Biol., Univ. Rome "Tor Vergata", Via della Ricera Scientifica 00133 Rome, Italy
SO AIDS Research and Human Retroviruses 12 (6). 1996. 491-496. ISSN: 0889-2229
LA English
AB This article reports the HIV epitope specificity of antibodies present in the sera of HIV-negative patients with autoimmune diseases. Recombinant gp120 and a panel of synthetic peptides derived from the amino acid consensus **sequences** of either related (gp120, gp41, and p24) or unrelated (**Mage-1**, necdin, heat shock protein (65 kDa), and amyloid) HIV proteins were tested by a specific ELISA. The first set of experiments performed on four patients with Sjogren's syndrome (SjS) and four patients with systemic lupus erythematosus (SLE) revealed a significant anti-gp120 antibody reactivity in autoimmune patients when compared to healthy HIV-negative controls. Moreover, such binding could be almost completely inhibited by preincubation with free gp120. A significant anti-p24 reactivity was observed in 18 of 29 sera from SjS patients and in 13 of 25 sera from SLE patients, while anti-gp41 was observed only in 3 of 14 SjS and in 2 of 20 SLE-affected patients. Similar analyses were performed in the murine model of autoimmunity, showing that sera from MRL/lpr mice were able to bind all HIV-related peptides in an age-dependent manner. The analysis of a panel of HIV unrelated peptides showed that SLE as well as MRL/lpr sera bind both HIV-related and unrelated peptides, while SjS sera failed to do so, revealing the polygonal nature of the SLE and MRL/lpr repertoire and the oligoclonal reactivity of SjS sera. This is also supported by inhibition experiments, which showed that SLE, but not SjS, sera competitively inhibited the binding to HIV gp120 peptide of sera from autoimmune MRL/lpr mice. These results indicate that an overlapping polyclonal repertoire is present in both SLE and MRL/lpr sera, while the oligoclonal specificity of SjS antibodies may be related to a specific, nonpolyclonal, activation against putative retroviral antigens.

L5 ANSWER 7 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:257365 BIOSIS
DN 98813494
TI Immunization of melanoma patients with melanoma cell vaccine induces anti-**MAGE-1** immunity.
AU Okamoto T; Yuzuki D; Morton D L; Hoon D S B
CS John Wayne Cancer Inst., Saint John's Hosp., Santa Monica, CA, USA
SO 87th Annual Meeting of the American Association for Cancer Research, Washington, D.C., USA, April 20-24, 1996. Proceedings of the American Association for Cancer Research Annual Meeting 37 (0). 1996.

478-479. ISSN: 0197-016X
DT Conference
LA English

L5 ANSWER 8 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:110455 BIOSIS
DN 98682590
TI **MAGE-1**-specific precursor cytotoxic T-lymphocytes present among tumor-infiltrating lymphocytes from a patient with breast cancer: Characterization and antigen-specific activation.
AU Toso J F; Oei C; Oshidari F; Tartaglia J; Paoletti E; Lyerly H K; Talib S; Weinhold K J
CS Dep. Surg., Box 2926, Duke Univ. Med. Cent., Durham, NC 27710-2926, USA
SO Cancer Research 56 (1). 1996. 16-20. ISSN: 0008-5472
LA English
AB A potential target for development of tumor-specific immunotherapeutic strategies is the **MAGE-1** gene. We have utilized a recently developed recombinant canarypox (ALVAC) virus vector containing the **MAGE-1** gene (vCP235) to activate CTLs from a breast cancer patient bearing a **MAGE -1**+ tumor. Tumor-infiltrating lymphocytes (TILs) obtained from the tumor of a patient were stimulated in vitro with irradiated autologous peripheral blood mononuclear cells acutely infected with the vCP235 construct. These TILs preferentially expanded approximately 6-fold over a 16-day culture period and specifically recognized an allogeneic transformed B-cell line acutely infected with a vaccinia-**MAGE-1** recombinant targeting vector (vP1188) in the context of HLA-A2 and/or B7. TCR V-beta analysis of in vitro expanded T cells by a quantitative multiprobe RNase protection assay revealed preferential expansion of TCR V-beta-6.3 and V-beta-6.4. In addition, homologous T-cell receptor beta CDR3 joining **sequences** were found in the in vitro stimulated cultures. These results suggest that tumor antigen-specific, MHC-restricted CTLs may be derived from precursor CTLs present in TILs obtained from patients with **MAGE-1**+ tumors by in vitro stimulation with recombinant avipox **MAGE-1** virus-infected autologous cells. Collectively, these findings provide a rationale for tumor-associated antigen-based immunization as a means of activating precursor CTLs residing in patients with tumors expressing defined tumor-associated antigens such as **MAGE-1**.

L5 ANSWER 9 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:63419 BIOSIS
DN 98635554
TI Human neoplasms elicit multiple specific immune responses in the autologous host.
AU Sahin U; Tuereci O; Schmitt H; Cochlovius B; Johannes T; Schmits R; Luo F S G; Schobert I; Pfreundschuh M
CS Med. Klinik Poliklinik, Innere Med. I, Univ. Saarlandes, D-66421 Homburg, Germany
SO Proceedings of the National Academy of Sciences of the United States of America 92 (25). 1995. 11810-11813. ISSN: 0027-8424
LA English
AB Expression of cDNA libraries from human melanoma, renal cancer, astrocytoma, and Hodgkin disease in Escherichia coli and screening

for clones reactive with high-titer IgG antibodies in autologous patient serum lead to the discovery of at least four antigens with a restricted expression pattern in each tumor. Besides antigens known to elicit T-cell responses, such as **MAGE-1** and tyrosinase, numerous additional antigens that were overexpressed or specifically expressed in tumors of the same type were identified. **Sequence** analyses suggest that many of these molecules, besides being the target of a specific immune response, might be of relevance for tumor growth. Antibodies to a given antigen were usually confined to patients with the same tumor type. The unexpected frequency of human tumor antigens, which can be readily defined at the molecular level by the serological analysis of autologous tumor cDNA expression cloning, indicates that human neoplasms elicit multiple specific immune responses in the autologous host and provides diagnostic and therapeutic approaches to human cancer.

L5 ANSWER 10 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:551288 BIOSIS
DN 98565588
TI Expression of the MAGE gene family in human head-and-neck squamous-cell carcinomas.
AU Eura M; Ogi K; Chikamatsu K; Lee K D; Nakano K; Masuyama K; Itoh K; Ishikawa T
CS Dep. Otolaryngol., Kumamoto Univ. Sch. Med., Honjo 1-1-1, Kumamoto 860, Japan
SO International Journal of Cancer 64 (5). 1995. 304-308. ISSN: 0020-7136
LA English
AB The MAGE genes encode certain tumor-associated antigens recognized by cytotoxic T lymphocytes. We investigated the expression of the **MAGE-1**, -2, -3, -4, -41 and -6 genes in 88 head-and-neck squamous-cell carcinomas (83 fresh tumor samples and 5 cell lines), using a reverse-transcription-polymerase-chain-reaction assay, followed by dot-blot hybridization with **sequence** -specific oligonucleotides and/or restriction-enzyme-pattern analysis. The **MAGE-1**, -2, -3, -4, -41 and -6 genes were expressed at the mRNA level in 27, 34, 36, 22, 16 and 35, respectively, of 83 fresh tumor samples. At least one of these genes was expressed in 59 of the 83 samples. Neither non-tumor inflammatory cells nor normal tissues were positive for these genes. The **MAGE-1** gene was expressed relatively frequently in SCC of the oropharynx, hypopharynx and maxillary sinus, but at lower rates in SCC of the larynx and of the tongue and oral cavity. **MAGE-1** was frequently expressed in poorly differentiated SCC, somewhat less frequently in moderately differentiated SCC, and only infrequently in well-differentiated SCC. The expression levels of the other MAGE genes also varied with the anatomic site as well as the degree of differentiation. Our results suggest that specific immunotherapy against MAGE gene products may be useful for patients with head-and-neck carcinomas.

L5 ANSWER 11 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:487219 BIOSIS
DN 98501519
TI Multiple specificities in the repertoire of a melanoma patient's cytolytic T lymphocytes directed against tumor antigen **MAGE** -1.A1.

AU Romero P; Pannetier C; Herman J; Jongeneel C V; Cerottini J-C; Coulie P G
CS Ludwig Inst. Cancer Res., Ch. des Boveresses 155, CH-1066 Epalinges, Switzerland
SO Journal of Experimental Medicine 182 (4). 1995. 1019-1028. ISSN: 0022-1007
LA English
AB Peptide **MAGE-1**.A1 is a nonamer derived from protein **MAGE-1** that can associate with the HLA-A1 molecule. It was shown previously to be recognized by an antitumor cytolytic T lymphocyte (CTL) clone derived from the blood of melanoma patient MZ2. We derived two other anti-**MAGE-1**.A1 CTL clones from different blood samples of the same patient and compared the fine specificity of recognition of the three CTL by testing them on variant **MAGE-1**.A1 peptides incorporating different amino acid substitutions. The epitopes recognized by the CTL proved to be different. While modifications of residues at positions 5, 6, or 7 in the antigenic peptide affected recognition by the three CTL, each of the modifications of residues at positions 1, 4, or 8 affected recognition by one CTL only. The **sequences** of both the alpha and beta chains of the T cell antigen receptor of the three CTL were completely different. The results indicate a long-lasting diversity in terms of fine specificity and of T cell antigen receptor structure in the repertoire of antitumor CTL derived From the blood of a melanoma patient and directed against a defined tumor antigen.

L5 ANSWER 12 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:440692 BIOSIS
DN 98454992
TI **Sequence** analysis of the MAGE gene family encoding human tumor-rejection antigens.
AU Imai Y; Shichijo S; Yamada A; Katayama T; Yano H; Itoh K
CS Dep. Immunol., Kurume Univ. Sch. Med., Kurume 830, Japan
SO Gene (Amsterdam) 160 (2). 1995. 287-290. ISSN: 0378-1119
LA English
AB The MAGE multigene family, which includes the **MAGE-1** and -3 genes that encode tumor-rejection antigens on HLA-A1 recognized by cytotoxic T-lymphocytes (CTL), is preferentially expressed at the mRNA level on human malignant cells, but not on normal cells. However, little is known about the MAGE-4, -41 and -6 genes. In this study, we have amplified 1040 hp (**MAGE-1**), 1061 hp (MAGE-3 and-6) and 1064 hp (MAGE-4 and -41) cDNA fragments, including the entire coding **sequences** (927-951 bp), using the reverse transcription-polymerase chain reaction (RT-PCR) method followed by nucleotide (nt) sequencing. One member had greater than 80 or 66% homology with the other members at the nt or deduced amino acid (aa) levels, respectively. Higher homology was found between MAGE-3 and -6 (98% at the nt level) and also between MAGE-4 and -41 (98%). The results of this investigation demonstrated high homology, as well as the clear differences between the members of the MAGE family at the coding **sequence** level.

L5 ANSWER 13 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:317508 BIOSIS
DN 98331808
TI Presentation of synthetic peptide antigen encoded by the **MAGE**

-1 gene by granulocyte-macrophage-colony-stimulating-factor-cultured macrophages from HLA-A1 melanoma patients.
AU Yamasaki S; Okino T; Chakraborty N G; Adkisson W O; Sampieri A; Padula S J; Mauri F; Mukherji B
CS Dep. Med., Univ. Conn. Health Cent., Farmington, CT 06030-3210, USA
SO Cancer Immunology Immunotherapy 40 (4). 1995. 268-271. ISSN: 0340-7004
LA English
AB The recent identification of the **sequences** of the peptides derived from a number of human melanoma-associated antigens has presented opportunities for developing a specific-peptide-based vaccine in this form of cancer. Since antigen-presenting cells (APC) play a crucial role in the induction of the T-cell-mediated immune response, we examined whether or not ex vivo cultured APC, bearing the appropriate MHC restricting elements, when pulsed with a relevant melanoma-specific cytotoxic-T-lymphocyte(CTL)-determined peptide, can present the peptide to the CTL. Here we show that a population of cells, derived from the monocyte/macrophage lineage from peripheral blood and grown in granulocyte/macrophage-colony-stimulating factor, exhibit many essential characteristics of "professional" APC (dendritic-type morphology with a proportion of the population, the B7 molecule, and high levels of MHC class I and class II molecules, CD11b and CD54 molecules) and are capable of efficiently presenting the nonapeptide, EADPTGHSY, encoded by the melanoma antigen **MAGE-1** gene, to the **MAGE-1** -specific CTL clone, 82/30. These results suggest that this type of autologous ex vivo cultured population of professional APC, when pulsed with the relevant-CTL-determined peptide, can serve as a novel type of candidate vaccine for active specific immunization against HLA-A1-positive patients with melanoma expressing the **MAGE-1** antigen.

L5 ANSWER 14 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 95:128084 BIOSIS
DN 98142384
TI Identification of potential CTL epitopes of tumor-associated antigen **MAGE-1** for five common HLA-A alleles.
AU Celis E; Fikes J; Wentworth P; Sidney J; Southwood S; Maewal A; Del Guercio M-F; Sette A; Livingston B
CS 3525 John Hopkins Court, Cytel Corp., San Diego, CA 92121, USA
SO Molecular Immunology 31 (18). 1994. 1423-1430. ISSN: 0161-5890
LA English
AB Identification of CTL epitopes for tumor-specific responses is important for the development of immunotherapies to treat cancer patients. We have developed a strategy to identify potential CTL epitopes based on screening of **sequences** of target proteins for presence of specific motifs recognized by the most common HLA-A alleles, and identification of high affinity binding peptides using in vitro quantitative assays. A systematic analysis using the **sequence** of the product of the tumor-associated **MAGE** -1 gene has been carried out. All possible peptides of nine and ten residues, containing binding motifs for HLA-A1, -A2.1, A-3.2, -A11 and -A24 were synthesized and tested for binding using a quantitative assay. Out of 237 possible peptide/MHC combinations, 47 cases demonstrated good binding affinity (K-d ltoreq 500 nM). Several peptides were identified as good MHC binders for each one of the five HLA-A alleles studied (five for HLA-A1, 11 for HLA-A2.1, 10 for

HLA-A3.2, 16 for HLA-A11 and five for HLA-A24. Furthermore, eight of these peptides were found to bind well to more than one HLA-A allele. These results have important implications for the development of immunotherapeutic vaccines to treat malignant melanoma.

L5 ANSWER 15 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 94:545758 BIOSIS
DN 98005306
TI Structure, chromosomal localization, and expression of 12 genes of the MAGE family.
AU De Plaen E; Arden K; Traversari C; Gaforio J J; Szikora J-P; De Smet C; Brasseur F; Van Der Bruggen P; Lethe B; Lurquin C; Brasseur R; Chomez P; De Backer O; Cavenee W; Boon T
CS Ludwig Inst. Cancer Res., Brussels Branch, 74 Ave. Hippocrate, B-1200 Brussels, Belgium
SO Immunogenetics 40 (5). 1994. 360-369. ISSN: 0093-7711
LA English
AB We reported previously that human gene **MAGE-1** directs the expression of a tumor antigen recognized on a melanoma by autologous cytolytic T lymphocytes. Probing cosmid libraries with a **MAGE-1 sequence**, we identified 11 closely related genes. The analysis of hamster-human somatic cell hybrids indicated that the 12 MAGE genes are located in the q terminal region of chromosome X. Like **MAGE-1**, the 11 additional MAGE genes have their entire coding **sequence** located in the last exon, which shows 64%-85% identity with that of MAGE1. The coding **sequences** of the MAGE genes predict the same main structural features for all MAGE proteins. In contrast, the promoters and first exons of the 12 MAGE genes show considerable variability, suggesting that the existence of this gene family enables the same function to be expressed under different transcriptional controls. The expression of each MAGE gene was evaluated by reverse transcription and polymerase chain reaction amplification. Six genes of the MAGE family including **MAGE-1** were found to be expressed at a high level in a number of tumors of various histological types. None was expressed in a large panel of healthy tissues, with the exception of testis and placenta.

L5 ANSWER 16 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 94:393049 BIOSIS
DN 97406049
TI Cloning and analysis of **MAGE-1**-related genes.
AU Ding M; Beck R J; Keller C J; Fenton R G
CS NCI-FCRDC, P.O. Box B, Bldg. 567, Room 207, Frederick, MD 21702, USA
SO Biochemical and Biophysical Research Communications 202 (1). 1994. 549-555. ISSN: 0006-291X
LA English
AB The spectrum of MAGE gene expression in the human melanoma cell line DM150 was examined using reverse transcription polymerase chain reaction and cDNA cloning. We have isolated five full-length cDNAs from DM150 which were identified as **MAGE-1**, MAGE-3, MAGE-12 and two previously undescribed MAGE genes, MAGE-3b and MAGE-X2. DNA **sequence** analysis of the coding regions of the MAGE-3b and MAGE-X2 genes revealed 83% and 88% identity with **MAGE-1**, while MAGE-3b was 98% homologous with the full length MAGE-3 clone. The predicted amino acid **sequences** of MAGE-X2 and MAGE-3b contain consensus HLA-A1 peptide binding

motifs, suggesting that, like **MAGE-1**, they may code for tumor-associated antigens. In addition, a nonamer peptide encoded by both the MAGE-3 and MAGE-12 genes was shown by direct binding studies to contain an aggretope for HLA-A2.

L5 ANSWER 17 OF 17 BIOSIS COPYRIGHT 1998 BIOSIS
AN 94:180223 BIOSIS
DN 97193223
TI Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous cytolytic T lymphocytes.
AU Gaugler B; Van Den Eynde B; Van Der Bruggen P; Romero P; Gaforio J J; De Plaen E; Lethe B; Brasseur F; Boon T
CS Ludwig Institute for Cancer Research, Brussels Branch, 74 Ave. Hippocrate, UCL 74-59, B-1200 Brussels, BEL
SO Journal of Experimental Medicine 179 (3). 1994. 921-930. ISSN: 0022-1007
LA English
AB Human melanoma cell line MZ2-MEL expresses several antigens recognized by autologous cytolytic T lymphocyte (CTL) clones. We reported previously the identification of a gene, named **MAGE -1**, that codes for one of these antigens named MZ2-E. We show here that antigen MZ2-D, which is present on the same tumor, is encoded by another member of the MAGE gene family named MAGE-3. Like **MAGE-1**, MAGE-3 is composed of three exons and the large open reading frame is entirely located in the third exon. Its **sequence** shows 73% identity with **MAGE-1**. Like MZ2-E, antigen MZ2-D is presented by HLA-A1. The antigenic peptide of MZ2-D is a nonapeptide that is encoded by the **sequence** of MAGE-3 that is homologous to the **MAGE-1 sequence** coding for the MZ2-E peptide. Competition experiments using single Ala-substituted-peptides indicated that amino acid residues Asp in position 3 and Tyr in position 9 were essential for binding of the **MAGE-1** peptide to HLA-A1. Gene MAGE-3 is expressed in many tumors of several types, such as melanoma, head and neck squamous cell carcinoma, lung carcinoma and breast carcinoma, but not in normal tissues except for testes. It is expressed in a larger proportion of melanoma samples than **MAGE-1**. MAGE-3 encoded antigens may therefore have a wide applicability for specific immunotherapy of melanoma patients.

published march 1994 MC

L12 ANSWER 1 OF 8 BIOSIS COPYRIGHT 1998 BIOSIS
AN 98:168613 BIOSIS
DN 01168613
TI Identification of a human VPF-VEGF 3' untranslated region mediating hypoxia-induced mRNA stability.
AU Claffey K P; Shih S-C; Mullen A; Dziennis S; Cusick J L; Abrams K R; Lee S W; Detmar M
CS Dep. Pathol., Beth Israel Deaconess Med. Cent., Boston, MA 02215, USA
SO Molecular Biology of the Cell 9 (2). 1998. 469-481. ISSN: 1059-1524
LA English
AB Hypoxia is a prominent feature of malignant tumors that are characterized by angiogenesis and vascular hyperpermeability. Vascular permeability factor/vascular endothelial growth factor (VPF/VEGF) has been shown to be up-regulated in the vicinity of

necrotic tumor areas, and hypoxia potently induces VPF/VEGF expression in several tumor cell lines in vitro. Here we report that hypoxia-induced VPF/VEGF expression is mediated by increased transcription and mRNA stability in human M21 **melanoma** cells. RNA-binding/electrophoretic mobility shift assays identified a single 125-bp AU-rich element in the 3' untranslated region that formed hypoxia-inducible RNA-protein complexes. Hypoxia-induced expression of chimeric luciferase reporter constructs containing this 125-bp AU-rich hypoxia stability region were significantly higher than constructs containing an adjacent 3' untranslated region element without RNA-binding activity. Using UV-cross-linking studies, we have identified a series of hypoxia-induced proteins of 90/88 kDa, 72 kDa, 60 kDa, 56 kDa, and **46 kDa** that bound to the hypoxia stability region element. The 90/88-kDa and 60-kDa species were specifically competed by excess hypoxia stability region RNA. Thus, increased VPF/VEGF mRNA stability induced by hypoxia is mediated, at least in part, by specific interactions between a defined mRNA stability sequence in the 3' untranslated region and distinct mRNA-binding **proteins** in human **tumor** cells.

ST RESEARCH ARTICLE; HUMAN; HYPOXIA; MESSENGER RNA STABILITY; VASCULAR PERMEABILITY FACTOR; VASCULAR ENDOTHELIAL GROWTH FACTOR; MICROVASCULAR PERMEABILITY; ANGIOGENESIS; VASCULAR HYPERPERMEABILITY; TUMOR; TUMOR BIOLOGY; NEOPLASTIC DISEASE

L12 ANSWER 2 OF 8 BIOSIS COPYRIGHT 1998 BIOSIS
AN 97:111308 BIOSIS
DN 99410511
TI The expression of the secreted protein acidic and rich in cystein (SPARC) is associated with the neoplastic progression of human **melanoma**.
AU Ledda F; Bravo A I; Adris S; Bover L; Mordoh J; Podhajcer O L
CS Instituto de Investigaciones Bioquimicas Luis F. Leloir "Fundacion Campomar", Av. Patricias Argentinas 435, 1405 Buenos Aires, Argentina
SO Journal of Investigative Dermatology 108 (2). 1997. 210-214. ISSN: 0022-202X
LA English
AB SPARC (secreted protein acidic and rich in cysteine) is an extracellular protein associated with tissues exhibiting high rates of cell proliferation and matrix remodeling. The current work shows that the human **melanoma** cell lines IIB-MEL-LES, IIB-MEL-IAN, and IIB-MEL-J and different human metastatic **melanomas** expressed high levels of SPARC mRNA and protein. By western blot analysis we detected a single secreted 42-kDa band in human diploid fibroblasts-conditioned medium and a 45-to 40-kDa doublet in the three **melanoma** cell lines and all the metastatic **melanomas** tested. Part of the **melanoma** samples and cell lines showed an additional doublet of 36-**34 kDa**. SPARC mRNA was expressed by the three established cell lines, 14 metastatic **melanoma** samples, and tumors raised in nude mice, and no spliced variants were found. The heterogeneous pattern of SPARC secreted by human **melanoma** cells is the result of post-translational glycosylation and a specific extracellular leupeptin-inhibitable cleavage. Unlike human fibroblasts, **melanoma** cells did not overexpress SPARC on addition of TGF-beta. Immunohistochemical analysis showed that SPARC was strongly expressed in 100% of primary **melanomas** (7 of

7) and metastatic **melanomas** (29 of 29), moderately expressed in most of the positive dysplastic nevi (13 of 14), and only weakly expressed in nevocellular nevi (4 of 25). Normal melanocytes did not express SPARC. The data suggest that the expression of SPARC is associated with the neoplastic progression of human **melanoma**.

ST RESEARCH ARTICLE; NUDE MOUSE; IIB-MEL-LES CELL LINE; IIB-MEL-IAN CELL LINE; IIB-MEL-J CELL LINE; HUMAN **MELANOMA** CELLS; **MELANOMA**; INTEGUMENTARY SYSTEM; **TUMOR** BIOLOGY; SECRETED **PROTEIN** ACIDIC AND RICH IN CYSTEINE; EXPRESSION; DERMAL FIBROBLAST; IMMUNOHISTOCHEMISTRY; NEVUS; BIOCHEMISTRY AND BIOPHYSICS; METASTASES; NEOPLASTIC DISEASE; INTEGUMENTARY SYSTEM; IMMUNOLOGICAL METHOD; DYSPLASTIC; INTEGUMENTARY SYSTEM DISEASE; NEVOCELLULAR

L12 ANSWER 3 OF 8 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:415356 BIOSIS
DN 99137712
TI Monoclonal antibodies against recombinant-**MAGE-1** protein identify a cross-reacting 72-kDa antigen which is co-expressed with **MAGE-1** protein in **melanoma** cells.
AU Carrel S; Schreyer M; Spagnoli G; Cerottini J-C; Rimoldi D
CS Ludwig Inst. Cancer Res., Ch. des Boveresses 155, 1066 Epalinges, Switzerland
SO International Journal of Cancer 67 (3). 1996. 417-422. ISSN: 0020-7136
LA English
AB The MAGE-I gene codes for tumor-associated peptides recognized by cytolytic T lymphocytes in association with MHC-class-I molecules such as HLA-A I and HLA-Cw16. In the course of a study aiming at the immunohistochemical detection of the MAGE-I gene product in tumor samples, 2 mouse monoclonal antibodies (MAbs) directed against a full-length recombinant MAGE-I fusion protein were found to react strongly not only with the **46-kDa** MAGE-I protein, but also with a 72-kDa product in immunoblots of lysates obtained from several MAGE1-mRNA-positive **melanoma** cell lines. Pre-incubation of the antibodies with the recombinant MAGE-I fusion protein abolished their reactivity both with MAGE-I protein and with the 72-kDa product, thus confirming the occurrence of antigenic determinant(s) shared by the 2 proteins. The 72-kDa protein is not an alternative product of MAGE-I, since it was still detected in lysates of a MAGE-I loss variant derived from a MAGE-I-positive **melanoma** cell line. Moreover, the 72-kDa protein does not appear to be a product of the other members of the MAGE gene family known to be expressed in tumors (such as MAGE-2, -3, -4 and -12). Interestingly, expression of the 72-kDa protein was found to be correlated with that of MAGE-I **protein**. Thus, in 30 **tumor** cell lines analyzed by immunoblotting and RT-PCR, the 72-kDa protein was never detected in MAGE-I-mRNA-negative cell lines, while it was co-expressed with MAGE-I protein in 12 out of 15 cell lines expressing MAGE-I. Furthermore, the 72-kDa protein was detected in lysates of human testis, the only normal tissue known to express MAGE-I. Finally, treatment of **MAGE- 1** -mRNA-negative cell lines with 5-Aza-2'-deoxycytidine, a hypomethylating agent known to induce MAGE- I expression, resulted in the expression of the 72-kDa protein. Taken collectively, these

findings suggest that expression of the gene encoding the 72-kDa protein identified in this study through antigenic determinant(s) shared with MAGE-I protein is regulated in a way similar to that of MAGE-I.
ST RESEARCH ARTICLE; HUMAN; MOUSE; MESSENGER RNA; HLA ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX; CYTOLYTIC T-LYMPHOCYTES; 5-AZA-2'-DEOXYCYTIDINE; METABOLIC-DRUG

L12 ANSWER 4 OF 8 BIOSIS COPYRIGHT 1998 BIOSIS
AN 96:110561 BIOSIS
DN 98682696
TI Expression of **MAGE-1**, MAGE-2, MAGE-3--6, and MAGE-4a--4b genes in ovarian tumors.
AU Yamada A; Kataoka A; Shichijo S; Kamura T; Imai Y; Nishida T; Itoh K
CS Dep. Immunol., Kurume Univ. Sch. Med., 67 Asahi-machi, Kurume 830, Japan
SO International Journal of Cancer 64 (6). 1995. 388-393. ISSN: 0020-7136
LA English
AB MAGE genes encoding tumor-rejection antigens recognized by cytotoxic T lymphocytes are expressed at the mRNA level in various malignant tumors. We have investigated the expression of genes **MAGE-1**, -2, -3/-6 and -4a/-4b at the mRNA level in malignant and non-malignant ovarian tumors as well as in normal ovaries by reverse transcription-polymerase chain reaction. **MAGE-1**, -2. -3/-6 and -4a/-4b were expressed in 12, 5, 11 and 4 of 58 malignant tumors, respectively- The majority of these MAGE-mRNA-positive tumors were histologically surface-epithelial-stromal tumors, in particular serous adenocarcinomas. They mostly consisted of either advanced-stage or recurrent tumors. In contrast, neither benign tumors nor normal ovaries expressed any of the MAGE genes investigated. A **46-kDa MAGE-1** protein was identified in **MAGE-1** -mRNA-positive serous adenocarcinomas by immunoblot analysis with polyclonal anti-**MAGE-1** antibody. These results provide important information for specific immunotherapy of ovarian serous e products.
ST RESEARCH ARTICLE; HUMAN; MESSENGER RNA; STAGING; ADENOCARCINOMA

L12 ANSWER 5 OF 8 BIOSIS COPYRIGHT 1998 BIOSIS
AN 94:548546 BIOSIS
DN 98008094
TI **MAGE-1** gene product is a cytoplasmic protein.
AU Schultz-Thater E; Juretic A; Dellabona P; Luscher U; Siegrist W; Harder F; Heberer M; Zuber M; Spagnoli G C
CS Z.L.F., Surgical Res. Lab., 20 Hebelstrasse, CH-4031 Basel, Switzerland
SO International Journal of Cancer 59 (3). 1994. 435-439. ISSN: 0020-7136
LA English
AB **MAGE-1** gene encodes a human **melanoma** antigen, recognized by syngeneic cytotoxic T lymphocytes (CTL). **MAGE-1** transcripts are also detectable in breast cancers, in non-small-cell lung carcinomas and in central nervous system tumors. In order to identify, in cellular preparations, the protein encompassing the antigenic peptide, we generated a panel of monoclonal antibodies (MAbs) against the **MAGE-1**

gene product by using, as immunogen, a full-length recombinant preparation (rMAGE1), obtained through expression cloning of the relevant gene in E. coli. Four reagents were obtained recognizing both rMAGE-1 and the **46-kDa** native protein in cell lines expressing **MAGE-1** mRNA. No positivity could be detected in **MAGE-1**-mRNA-negative **melanoma** lines. No surface labelling of **MAGE-1**-positive cell lines could be observed. In contrast, on permeabilization of MZ2 **melanoma** cells, all 4 MAbs induced efficient staining, as detected by cytofluorography. Fluorescence microscopy shows that **MAGE-1** gene product is a cytoplasmic protein clustered in paranuclear organelle-like structures. Thus, **MAGE-1** protein location closely resembles that of P91A and P198 murine-**tumor antigens**.

ST RESEARCH ARTICLE; ESCHERICHIA COLI; HUMAN; MOUSE; **MELANOMA** ANTIGEN; BREAST; LUNG; CENTRAL NERVOUS SYSTEM TUMORS; T-LYMPHOCYTE RECOGNITION; MESSENGER RNA; MZ-2 CELL LINE

L12 ANSWER 6 OF 8 BIOSIS COPYRIGHT 1998 BIOSIS
AN 94:303423 BIOSIS
DN 97316423
TI Characterization of cluster 13: The epithelial-carcinoma antigen recognized by MAb RS7.
AU Stein R; Basu A; Goldenberg D M; Lloyd K O; Mattes M J
CS Center Mol. Med. Immunol., One Bruce St., Newark, NJ 07103, USA
SO International Journal of Cancer 0 (SUPPL. 8). 1994. 98-102. ISSN: 0020-7136
LA English
AB Cluster 13 was defined by 2 independently derived murine monoclonal antibodies (MAbs), RS7 (IgG-1) and MR54 (IgG-2a), which were raised against human squamous-cell carcinoma of the lung and a human ovarian-carcinoma cell line, respectively. Immunologic and biochemical evidence demonstrated that RS7 and MR54, as well as 2 additional MAbs, MR6 (IgG-2a) and MR23 (IgG-1), generated in the same fusion as MR54, recognize the same antigen, a **46**- to 48-**kDa** glycoprotein. Evaluation of the expression of **antigen** on the surface of **tumor** cell lines, Western blotting analyses, competitive binding studies, and double-determinant ELISA assays, support this conclusion. Two distinct epitopes are defined by these MAbs. In order to further characterize this antigen, amino-acid-sequence analyses were performed on peptides derived from antigen purified by affinity chromatography with MAb RS7. The sequence data obtained from 2 peptides, which were independently generated by CNBr cleavage and trypsin digestion respectively indicated identity to GA733-1. The GA733-1 genomic DNA sequence predicted a type-1 membrane protein of 35 kDa, with 4 potential N-linked glycosylation sites. The GA733-1 protein product has not been identified previously, and MAbs to this **tumor**-associated **antigen** were not previously known.

ST RESEARCH ARTICLE; HUMAN; MONOCLONAL ANTIBODY; LUNG; OVARIAN CANCER CELLS; ELISA; IMMUNOLOGICAL METHOD; **MELANOMA**; COLON; PANCREATIC; CERVICAL CANCER CELLS; GLYCOPROTEIN

L12 ANSWER 7 OF 8 BIOSIS COPYRIGHT 1998 BIOSIS
AN 94:129901 BIOSIS
DN 97142901

TI Identification of the **MAGE-1** gene product by monoclonal and polyclonal antibodies.
AU Chen Y-T; Stockert E; Chen Y; Garin-Chesa P; Rettig W J; Van Der Bruggen P; Boon T; Old L J
CS Ludwig Inst. Cancer Res., New York Unit, New York Hosp.-Cornell Med. Cent., New York, NY 10021, USA
SO Proceedings of the National Academy of Sciences of the United States of America 91 (3). 1994. 1004-1008. ISSN: 0027-8424
LA English
AB The human **MAGE-1** gene encodes a **melanoma** peptide antigen recognized by autologous cytotoxic T lymphocytes. To produce antibodies against the **MAGE-1** gene product, several approaches were taken. Three oligopeptides were synthesized based on predicted **MAGE-1** amino acid sequences and were used to generate rabbit anti-peptide antisera. In addition, a truncated **MAGE-1** cDNA was cloned into an Escherichia coli expression vector, and recombinant protein was produced and purified. All three rabbit anti-peptide antisera showed reactivity against the immunizing peptide, and one reacted with the recombinant **MAGE-1** protein by immunoblotting, but none reacted with cell lysates from **MAGE-1** mRNA-positive cells. The recombinant **MAGE-1** protein was then used for the generation of mouse monoclonal and rabbit polyclonal antibodies. One IgG1 monoclonal antibody, MA454, as well as rabbit polyclonal antisera recognized a **46-kDa** protein in extracts of **MAGE-1** mRNA-positive **melanoma** cell lines. The antibodies showed no apparent crossreactivity with products of the closely related MAGE-2 and MAGE-3 genes. Serological typing of normal and tumor cell lysates was in full agreement with mRNA analysis, showing expression of **MAGE-1** protein in **MAGE-1** mRNA-positive testis and a subset of **melanomas** but not in **MAGE-1** mRNA-negative normal or tumor tissues. Transfection of the **MAGE-1** gene into a **MAGE-1** mRNA-negative **melanoma** cell line resulted in the expression of the **46-kDa** protein, confirming the identity of this protein as the **MAGE-1** gene product.
ST RESEARCH ARTICLE; HUMAN **MELANOMA** CELLS; RABBIT; MOUSE; **TUMOR**-REJECTION **ANTIGEN**

L12 ANSWER 8 OF 8 BIOSIS COPYRIGHT 1998 BIOSIS
AN 90:471686 BIOSIS
DN BA90:111106
TI MOLECULAR CLONING OF COMPLEMENTARY DNA FOR THE HUMAN **TUMOR**-ASSOCIATED **ANTIGEN** CO-029 AND IDENTIFICATION OF RELATED TRANSMEMBRANE ANTIGENS.
AU SZALA S; KASAI Y; STEPLEWSKI Z; RODECK U; KOPROWSKI H; LINNENBACH A J
CS WISTAR INST. ANAT. BIOL., 3601 SPRUCE ST., PHILADELPHIA, PA. 19104.
SO PROC NATL ACAD SCI U S A 87 (17). 1990. 6833-6837. CODEN: PNASA6 ISSN: 0027-8424
LA English
AB The human **tumor**-associated **antigen** CO-029 is a monoclonal antibody-defined cell surface glycoprotein of 27-**34 kDA**. By using the high-efficiency COS cell expression system, a full-length cDNA clone for CO-029 was isolated. When transiently expessed in COS cells, the cDNA clone directed the

synthesis of an antigen reactive to monoclonal antibody CO-029 in mixed hemadsorption and immunoblot assays. Sequence analysis revealed that CO-029 belongs to a family of cell surface antigens that includes the **melanoma**-associated antigen ME491, the leukocyte cell surface antigen CD37, and the Sm23 antigen of the parasitic helminth Schistosoma mansoni. CO-029 and ME491 antigen expression and the effect of their coresponding monoclonal antibodies on cell growth were compared in human tumor cell lines of various histologic origins.

ST SCHISTOSOMA-MANSONI SM23 ANTIGEN **MELANOMA**-ASSOCIATED ME491 ANTIGEN LEUKOCYTE CD37 ANTIGEN MONOCLONAL ANTIBODY MOLECULAR SEQUENCE DATA AMINO ACID SEQUENCE NUCLEOTIDE SEQUENCE GENBANK-M35252

=> fil wpids

FILE 'WPIDS' ENTERED AT 09:30:59 ON 07 APR 1998
COPYRIGHT (C) 1998 DERWENT INFORMATION LTD

FILE LAST UPDATED: 06 APR 1998 <19980406/UP>
>>>UPDATE WEEKS:
MOST RECENT DERWENT WEEK 199814 <199814/DW>
DERWENT WEEK FOR CHEMICAL CODING: 199809
DERWENT WEEK FOR POLYMER INDEXING: 199811
DERWENT WORLD PATENTS INDEX SUBSCRIBER FILE, COVERS 1963 TO DATE
>>> D COST AND SET NOTICE DO NOT REFLECT SUBSCRIBER DISCOUNTS - SEE HELP COST FOR DETAILS <<<
>>> MEXICO NOW COVERED - SEE NEWS <<<

=> d que

L13 18 SEA FILE=WPIDS ABB=ON (46 OR 34) (3W) KDA
L14 1 SEA FILE=WPIDS ABB=ON L13 AND (MAGE OR MELANOMA)

=> d bib ab

L14 ANSWER 1 OF 1 WPIDS COPYRIGHT 1998 DERWENT INFORMATION LTD
AN 93-035688 [04] WPIDS
DNC C93-016134
TI Vaccine against swine dysentery - contains Treponema hyodysenteriae haemolysin, opt. with immunogenic protein.
DC B04 C03 D16
IN MCCAMAN, M; SLOMIANY, R
PA (MLTE-N) ML TECHNOLOGY VENTURES LP
CYC 1
PI US 5176910 A 930105 (9304)* 4 pp
ADT US 5176910 A US 89-296958 890117
PRAI US 89-296958 890117
AB US 5176910 A UPAB: 931119
Vaccine for protection against swine dysentery comprises Treponema hyodysenterial hemolysin (I) free of T. hyodysenterial cells and a carrier. The vaccine pref. also contains at least one protein (II) which elicits antibodies which recognise a T. hyodysenterial antigen of mol. wt. 19-90 kda.

(II) elicits an antibody which recognises at least one of the 19 kda, 29 kda, 30 kda, 31 kda, **34 kda**, 36 **kda**, 38 kda, 39 kda, 42 kda, 44 kda, and 60 kdaT.

hyodysenterial antigens.
USE - In addn. to its use as a vaccine. (I) has been shown to act like an enterotoxin when applied to isolated swine mucosal tissue. (I) has also been shown to be cytopathic to CHO cells, Bowes **melanoma** cells and RBC
Dwg.0/0

```
=> fil hcaplus

FILE 'HCAPLUS' ENTERED AT 09:31:23 ON 07 APR 1998
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 1998 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is
held by the publishers listed in the PUBLISHER (PB) field (available
for records published or updated in Chemical Abstracts after December
26, 1996), unless otherwise indicated in the original publications.

FILE COVERS 1967 - 7 Apr 1998  VOL 128 ISS 15
FILE LAST UPDATED: 7 Apr 1998  (980407/ED)

 This file contains CAS Registry Numbers for easy and accurate
 substance identification.

 This file now supports REG1stRY for direct browsing and searching
 of all non-structural data from the REGISTRY file.  Enter HELP FIRST
 for more information.
'OBI' IS DEFAULT SEARCH FIELD FOR 'HCAPLUS' FILE

=> d his


     (FILE 'HCAPLUS' ENTERED AT 09:31:23 ON 07 APR 1998)
               DEL HIS Y
L1          163 S (46 OR 34) (L) KDA
L2         2536 S ((46 OR 34) (2W) KDA)/AB
L3         2563 S L1 OR L2
L4            4 S L3 AND MAGE
L5            4 S L3 AND (TUMOR# OR TUMOUR#) AND MELANOMA?
L6            8 S L4 OR L5

=> d .ca 1-8


L6   ANSWER 1 OF 8  HCAPLUS  COPYRIGHT 1998 ACS
AN   1997:710404  HCAPLUS
DN   127:357528
TI   Keratin 17: immunohistochemical mapping of its distribution in human
     epithelial tumors and its potential applications
AU   Miettinen, Markku; Nobel, Michael P.; Tuma, Bodil T.; Kovatich,
     Albert J.
CS   Department of Soft Tissue Pathology, Armed Forces Institute of
     Pathology, Washington, DC, 20306-6000, USA
SO   Appl. Immunohistochem. (1997), 5(3), 152-159
     CODEN: APIMEH; ISSN: 1062-3345
PB   Lippincott-Raven
```

DT Journal
LA English
AB The distribution of keratin 17 (K17), mol. wt. (MW) **46 kDa**, was immunohistochem. evaluated in a series of >500 formaldehyde-fixed, paraffin-embedded human epithelial and selected nonepithelial tumor specimens using a specific monoclonal antibody (clone E3). In normal epithelial tissues, K17 reactivity was found in a subset of hair shaft cells, sebaceous gland reserve cells, and respiratory and prostate basal epithelial cells. Furthermore, variable and inconsistent K17 expression was found in myoepithelia of sweat, salivary, and mammary glands. Although normal epidermis and other squamous epithelia were K17-neg., regenerative epithelia, e.g., at the site of ulceration or acute inflammation,. were pos. Furthermore, K17 was present in dysplastic but not in normal squamous epithelia of skin and internal organs. Most squamous-cell carcinomas showed K17-pos. cells that were more numerous in more well-differentiated tumors. About 50% of pulmonary and >50% of pancreatic adenocarcinomas showed K17-pos. cells. Although K17-pos. cells, in some cases, reflected adenosquamous differentiation, this was not the case in all K17-pos. adenocarcinomas. Approx. one third of mammary ductal carcinomas showed K17-pos. cells, whereas lobular carcinomas were invariably K17-neg. In situ lobular and ductal carcinomas often showed prominent K17-pos. myoepithelial cells; such cells showed even greater positivity when present with in situ carcinoma or proliferative conditions, as compared with normal ducts. Other adenocarcinomas often showing K17-pos. cells included endometrial, ovarian, and, to a lesser degree, gastric carcinomas. K17-neg. tumors, almost without exception, included colon, prostate, renal, and hepatocellular carcinomas and neuroendocrine tumors. Synovial sarcomas typically showed K17-pos. epithelial cells in biphasic but not in monophasic spindle cell tumors. The obsd. distribution of K17 indicates that this marker may be helpful in evaluating abnormalities of squamous epithelia and in the subtyping of epithelial tumors.
CC 14-1 (Mammalian Pathological Biochemistry)
ST keratin 17 epithelium **tumor**
IT Thyroid carcinoma
(anaplastic; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Prostate
Respiratory tract
(basal epithelium; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Carcinoid
(bronchial; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Squamous cell carcinoma
(cervical; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Digestive system **tumors**
(cholangioma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT **Tumors** (animal)
(chordoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Gastric **tumors**
(colon carcinoma-like; immunohistochem. mapping of distribution

keratin 17 in human epithelial **tumors**)
IT Breast carcinoma
(ductal, infiltrating; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Testicular **tumors**
(embryonal carcinoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Sarcoma
(epithelioid; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Squamous cell carcinoma
(esophageal; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Basal cell carcinoma
Colon adenocarcinoma
Dermatitis
Endometrial adenocarcinoma
Epidermis (skin)
Epithelium
Follicular carcinoma (thyroid)
Gastric adenocarcinoma
Hepatoma
Lung adenocarcinoma
Medullary thyroid carcinoma
Melanoma
Pancreatic adenocarcinoma
Papillary carcinoma (thyroid)
Prostatic adenocarcinoma
Renal cell carcinoma
Small-cell carcinoma (lung)
Squamous cell carcinoma
Squamous cell carcinoma (lung)
Squamous cell carcinoma (skin)
(immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Breast **tumors**
(infiltrating lobular; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Pancreatic **tumors**
(islet of Langerhans; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Keratins
RL: BOC (Biological occurrence); BIOL (Biological study); OCCU (Occurrence)
(keratin 17; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Squamous cell carcinoma
(larynx; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Mammary epithelium
Salivary gland
Sweat gland
(myoepithelium; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Neural **tumors**
(peripheral, sheath; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)

IT Salivary gland diseases
(pleomorphic adenoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Lung **tumors**
(pleura, mesothelioma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Sebaceous gland
(reserve cell; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Adenoma
(salivary gland pleomorphic; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT **Tumors** (animal)
(salivary gland, parotid adenoid cystic carcinoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Synovial membrane
(sarcoma, biphasic; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Synovial membrane
(sarcoma, monophasic spindle cell; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Testicular **tumors**
(seminoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Ovarian **tumors**
(serous cystadenoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Ovarian carcinoma
(serous papillary carcinoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Hair
(shaft cell; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Ulcer
(skin; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Intestinal **tumors**
(small intestine, carcinoid; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Intestinal **tumors**
(small intestine, colon carcinoma-like; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Larynx
(squamous cell carcinoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Bladder carcinoma
Cervical carcinoma
Esophageal carcinoma
(squamous cell; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Sarcoma
(synovial membrane, biphasic; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Sarcoma
(synovial membrane, monophasic spindle cell; immunohistochem. mapping of distribution keratin 17 in human epithelial

tumors)
IT Lymphoma
Thymus **tumors**
(thymoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Bladder carcinoma
(transition cell; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Pleural diseases
(**tumor**, mesothelioma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Small intestine
(**tumors**, carcinoid; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Biliary tract
(**tumors**, cholangioma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Small intestine
(**tumors**, colon carcinoma-like; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Salivary gland diseases
(**tumors**, parotid adenoid cystic carcinoma; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Islet of Langerhans
(**tumors**; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Skin diseases
(ulcer; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)
IT Ovarian carcinoma
(undifferentiated; immunohistochem. mapping of distribution keratin 17 in human epithelial **tumors**)

L6 ANSWER 2 OF 8 HCAPLUS COPYRIGHT 1998 ACS
AN 1996:517504 HCAPLUS
DN 125:192240
TI Monoclonal antibodies against recombinant **MAGE**-1 protein identify a cross-reacting 72-kDa antigen which is co-expressed with **MAGE**-1 protein in melanoma cells
AU Carrel, Stefan; Schreyer, Magali; Spagnoli, Giulio; Cerottini, Jean-Charles; Rimoldi, Donata
CS Ludwig Institute Cancer Research, University Lausanne, Epalinges, 1066, Switz.
SO Int. J. Cancer (1996), 67(3), 417-422
CODEN: IJCNAW; ISSN: 0020-7136
DT Journal
LA English
AB The MAGE-1 gene codes for tumor-assocd. peptides recognized by cytolytic T lymphocytes in assocn. with MHC class I mols. such as HLA-A1 and HLA-Cw16. In the course of a study aimed at the immunohistochem. detection of the MAGE-1 gene product in tumor samples, 2 mouse monoclonal antibodies (MAbs) directed against a full-length recombinant MAGE-1 fusion protein were found to react strongly not only with the **46-kDa** MAGE-1 protein, but also with a 72-kDa product in immunoblots of lysates obtained from several MAGE-1-mRNA-pos. melanoma cell lines.

Pre-incubation of the antibodies with the recombinant MAGE-1 fusion protein abolished their reactivity both with MAGE-1 protein and with the 72-kDa product, thus confirming the occurrence of antigenic determinant(s) shared by the 2 proteins. The 72-kDa protein is not an alternative product of MAGE-1, since it was still detected in lysates of a MAGE-1 loss variant derived from a MAGE-1-pos. melanoma cell line. Moreover, the 72-kDa protein does not appear to be a product of the other members of the MAGE gene family known to be expressed in tumors (such as MAGE-2, -3, -4 and -12). Interestingly, expression of the 72-kDa protein was correlated with that of MAGE-1 protein. Thus, in 30 tumor cell lines analyzed by immunoblotting and RT-PCR, the 72-kDa protein was never detected in MAGE-1-mRNA-neg. cell lines, while it was co-expressed with MAGE-1 protein in 12 out of 15 cell lines expressing MAGE-1. Furthermore, the 72-kDa protein was detected in lysates of human testis, the only normal tissue known to express MAGE-1. Finally, treatment of MAGE-1 mRNA-neg. cell lines with 5-aza-2'-deoxycytidine, a hypomethylating agent known to induce MAGE-1 expression, resulted in the expression of the 72-kDa protein. Taken collectively, these findings suggest that expression of the gene encoding the 72-kDa protein identified in this study through antigenic determinant(s) shared with MAGE-1 protein is regulated in a way similar to that of MAGE-1.

CC 14-1 (Mammalian Pathological Biochemistry)

IT Antigens
RL: BPR (Biological process); BIOL (Biological study); PROC (Process)
(**MAGE**-1; monoclonal antibodies to **MAGE**-1 cross-react with 72 kDa protein in melanoma cells)

IT Testis
(monoclonal antibodies to **MAGE**-1 cross-react with 72 kDa protein in)

IT Melanoma
(monoclonal antibodies to **MAGE**-1 cross-react with 72 kDa protein in melanoma cells)

IT Proteins, specific or class
RL: BOC (Biological occurrence); BPR (Biological process); MFM (Metabolic formation); PRP (Properties); BIOL (Biological study); FORM (Formation, nonpreparative); OCCU (Occurrence); PROC (Process)
(72,000-mol.-wt., monoclonal antibodies to **MAGE**-1 cross-react with 72 kDa protein in melanoma cells)

IT Antibodies
RL: BPR (Biological process); BIOL (Biological study); PROC (Process)
(monoclonal, monoclonal antibodies to **MAGE**-1 cross-react with 72 kDa protein in melanoma cells)

L6 ANSWER 3 OF 8 HCAPLUS COPYRIGHT 1998 ACS
AN 1996:150834 HCAPLUS
DN 124:228283
TI Expression of **MAGE**-1, **MAGE**-2, **MAGE**-3/-6 and **MAGE**-4a/-4b genes in ovarian tumors
AU Yamada, Akira; Kataoka, Akio; Shichijo, Shigeki; Kamura, Toshiharu; Imai, Yasuhisa; Nishida, Takashi; Itoh, Kyogo
CS School of Medicine, Kurume University, Kurume, 830, Japan
SO Int. J. Cancer (1995), Volume Date 1995, 64(6), 388-93
CODEN: IJCNAW; ISSN: 0020-7136
DT Journal

LA English
AB MAGE genes encoding tumor-rejection antigens recognized by cytotoxic T lymphocytes are expressed at the mRNA level in various malignant tumors. The authors have investigated the expression of genes MAGE-1, -2, -3/-6, and -4a/-4b at the mRNA level in malignant and non-malignant ovarian tumors as well as in normal ovaries by reverse transcription-polymerase chain reaction. MAGE-1, -2, -3/-6, and -4a/-4b were expressed in 12, 5, 11, and 4 of 58 malignant tumors, resp. The majority of these MAGE-mRNA-pos. tumors were histol. surface-epithelial-stromal tumors, in particular serous adenocarcinomas. They mostly consisted of either advanced-stage or recurrent tumors. In contrast, neither benign tumors nor normal ovaries expressed any of the MAGE genes investigated. A **46 -kDa** MAGE-1 protein was identified in MAGE-1-mRNA-pos. serous adenocarcinomas by immunoblot anal. with polyclonal anti-MAGE-1 antibody. These results provide important information for specific immunotherapy of ovarian serous adenocarcinomas with MAGE gene products.
CC 14-1 (Mammalian Pathological Biochemistry)
Section cross-reference(s): 3, 15
ST **MAGE** gene expression ovary tumor
IT Gene, animal
RL: BPR (Biological process); BIOL (Biological study); PROC (Process)
(**MAGE**-1; expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)
IT Gene, animal
RL: BPR (Biological process); BIOL (Biological study); PROC (Process)
(**MAGE**-2; expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)
IT Gene, animal
RL: BPR (Biological process); BIOL (Biological study); PROC (Process)
(**MAGE**-3; expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)
IT Gene, animal
RL: BPR (Biological process); BIOL (Biological study); PROC (Process)
(**MAGE**-4a; expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)
IT Gene, animal
RL: BPR (Biological process); BIOL (Biological study); PROC (Process)
(**MAGE**-4b; expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)
IT Gene, animal

RL: BPR (Biological process); BIOL (Biological study); PROC (Process)
(**MAGE**-6; expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Ovary, neoplasm
(expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Proteins, specific or class
RL: BOC (Biological occurrence); BIOL (Biological study); OCCU (Occurrence)
(gene **MAGE**-1; expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Ovary, neoplasm
(arrhenoblastoma, expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Ovary, neoplasm
(clear cell carcinoma, expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Ovary, neoplasm
(dysgerminoma, expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Ovary, neoplasm
(endometrioid carcinoma, expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Ovary, neoplasm
(granulosa cell, expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Ovary, neoplasm
(metastasis, expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Ovary, neoplasm
(mucinous adenocarcinoma, expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**-3/-6 genes in serous adenocarcinomas)

IT Mesonephros
(neoplasm, ovarian yolk sac tumor, expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE**

-3/-6 genes in serous adenocarcinomas)
IT Ovary, neoplasm
(serous adenocarcinoma, expression of **MAGE**-1, -2, -3/-6, and -4a/-4b genes in human ovarian tumors with the preferential expression of **MAGE**-1 and **MAGE** -3/-6 genes in serous adenocarcinomas)

L6 ANSWER 4 OF 8 HCAPLUS COPYRIGHT 1998 ACS
AN 1995:995009 HCAPLUS
DN 124:127118
TI Delipidated purified polysaccharidic compound D25 preparation, and injectable composition comprising same
IN Binz, Hans; Durand, Jacques; Cudennec, Claude-Alain; Normier, Gerard; Le Pape, Alain
PA Pierre Fabre Medicament, Fr.
SO PCT Int. Appl., 56 pp.
CODEN: PIXXD2
PI WO 9525128 A1 950921
DS W: AU, CA, JP, NZ, US
RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE
AI WO 95-FR311 950315
PRAI FR 94-3072 940316
DT Patent
LA French
AB A polysaccharide compd. D25, with a mol. wt. of .apprx.**34 kDa**, extd. from the membrane proteoglycans of Klebsiella pneumoniae, is delipidated and purified so that an aq. soln. contg. the compd. no longer contains any C16 fatty acids or peptides whether free or combined with the said compd. The use of said prepn. in imaging, given the targeting properties of said compd., is also disclosed. Thus an injectable formulation comprised the purified carbohydrate 1.0 mg in an aq. soln. contg. NaCl at 1.0 mg. Such a soln. was labeled with 99mTc by the SnF2 method with ascorbic acid as stabilizing agent and was used for imaging of melanoma.
IC ICM C08B037-00
ICS A61K051-00; A61K031-715
CC 63-6 (Pharmaceuticals)
IT **Melanoma**
(polysaccharidic formulation for imaging and diagnosis of)
IT Macrophage
(targeting properties of injectable formulation contg. polysaccharide for **tumor** imaging)

L6 ANSWER 5 OF 8 HCAPLUS COPYRIGHT 1998 ACS
AN 1995:222942 HCAPLUS
DN 122:262435
TI **MAGE**-1 gene product is a cytoplasmic protein
AU Schultz-Thater, Elke; Juretic, Antonio; Dellabona, Paolo; Luescher, Urs; Siegrist, Walter; Harder, Felix; Heberer, Michael; Zuber, Markus; Spagnoli, Giulio C.
CS Department of Surgery, University of Basel, Basel, Switz.
SO Int. J. Cancer (1994), 59(3), 435-9
CODEN: IJCNAW; ISSN: 0020-7136
DT Journal
LA English
AB MAGE-1 gene encodes a human melanoma antigen, recognized by syngeneic cytotoxic T lymphocytes (CTL). MAGE-1 transcripts are

also detectable in breast cancers, in non-small cell lung carcinomas and in central nervous system tumors. To identify, in cellular prepns., the protein encompassing the antigenic peptide, the authors generated a panel of monoclonal antibodies (MAbs) against the MAGE-1 gene product by using, as immunogen, a full-length recombinant prepn. (rMAGE-1), obtained through expression cloning of the relevant gene in E. coli. Four reagents were obtained recognizing both rMAGE-1 and the **46-kDa** native protein in cell lines expressing MAGE-1 mRNA. No positivity could be detected in MAGE-1-mRNA-neg. melanoma lines. No surface labeling of MAGE-1-pos. cell lines could be obsd. In contrast, on permeabilization of MZ2 melanoma cells, all 4 MAbs induced efficient staining, as detected by cytofluorog. Fluorescence microscopy shows that MAGE-1 gene product is a cytoplasmic protein clustered in paranuclear organelle-like structures. Thus, MAGE-1 protein location closely resembles that of P91A and P198 murine-tumor antigens.

CC 14-1 (Mammalian Pathological Biochemistry)
Section cross-reference(s): 15

IT Gene, animal
RL: BSU (Biological study, unclassified); BIOL (Biological study)
(**MAGE**-1; human **MAGE**-1 gene antigen is cytoplasmic protein in melanoma)

IT Antigens
RL: ANT (Analyte); BOC (Biological occurrence); ANST (Analytical study); BIOL (Biological study); OCCU (Occurrence)
(gene **MAGE**-1; human **MAGE**-1 gene antigen is cytoplasmic protein in melanoma)

IT Cytoplasm
Melanoma
(human **MAGE**-1 gene antigen is cytoplasmic protein in melanoma)

IT Antibodies
RL: BPN (Biosynthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(monoclonal, to gene **MAGE**-1 antigen; human **MAGE**-1 gene antigen is cytoplasmic protein in melanoma)

L6 ANSWER 6 OF 8 HCAPLUS COPYRIGHT 1998 ACS
AN 1994:131751 HCAPLUS
DN 120:131751
TI Identification of the **MAGE**-1 gene product by monoclonal and polyclonal antibodies
AU Chen, Yao Tseng; Stockert, Elisabeth; Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.; Van der Bruggen, P.; Boon, Thierry; Old, Lloyd J.
CS New York Unit, Ludwig Inst. Cancer Res., New York, NY, 10021, USA
SO Proc. Natl. Acad. Sci. U. S. A. (1994), 91(3), 1004-8
CODEN: PNASA6; ISSN: 0027-8424
DT Journal
LA English
AB The human MAG-1 gene encodes a melanoma peptide antigen recognized by autologous cytotoxic T lymphocytes. To produce antibodies against the MAGE-1 gene product, several approaches were taken. Three oligopeptides were synthesized based on predicted MAGE-1 amino acid sequences and were used to generate rabbit anti-peptide anti-sera. In addn., a truncated MAGE-1 cDNA was cloned into an

Escherichia coli expression vector, and recombinant protein was produced and purified. All three rabbit anti-peptide antisera showed reactivity against the immunizing peptide, and one reacted with the recombinant MAGE-1 protein by immunoblotting, but none reacted with cell lysates from MAGE-1 mNA-pos. cells. The recombinant MAGE-1 protein was then used for the generation of mouse monoclonal and rabbit polyclonal antibodies. One IgG1 monoclonal antibody, MA454, as well as rabbit polyclonal antisera recognized a **46-kDa** protein in exts. of MAGE-1 mRNA-pos. melanoma cell lines. The antibodies showed no apparent cross-reactivity with products of the closely related MAGE-2 and MAGE-3 genes. Serol. typing of normal and tumor cell lysates was in full agreement with mRNA anal., showing expression of MAGE-1 protein in MAGE-1 mRNA-pos. testis and a subset of melanomas but not in MAGE-1 mRNA-neg. normal or tumor tissues. Transfection of the MAGE-1 gene into a MAGE-1 mRNA-neg. melanoma cell line resulted in the expression of the **46-kDa** protein, confirming the identity of this protein as the MAGE-1 gene product.

CC 15-2 (Immunochemistry)

IT Proteins, specific or class
RL: PREP (Preparation)
(gene **MAGE**-1, monoclonal and polyclonal antibodies to, prepn. and reactivity of)

IT Antibodies
RL: PREP (Preparation)
(to **MAGE**-1 gene product, prepn. and reactivity of)

IT Antibodies
RL: PREP (Preparation)
(monoclonal, to **MAGE**-1 gene product, prepn. and reactivity of)

L6 ANSWER 7 OF 8 HCAPLUS COPYRIGHT 1998 ACS

AN 1994:125072 HCAPLUS

DN 120:125072

TI The melanocyte-stimulating hormone (MSH) receptor in M2R mouse **melanoma tumors**: solubilization and properties of the receptor-MSH complex and its covalently crosslinked conjugate

AU Shafir, I.; Schmidt-Sole, J.; Shai, E.; Salomon, Y.

CS Dep. Horm. Res., Weizmann Inst. Sci., Rehovot, 76100, Israel

SO Melanoma Res. (1993), 3(3), 157-68
CODEN: MREEEH; ISSN: 0960-8931

DT Journal

LA English

AB Several properties of the MSH receptor in solid melanotic and amelanotic mouse M2R tumor isografts were studied in C57BL mice. Using cell membrane fractions prepd. from such tumors and the superpotent [Nle4,D-Phe7].alpha.MSH analog, the affinity and receptor contents of the two tumor variants were similar. When occupied by MSH, the receptor-MSH complex (R.cntdot.MSH) was readily sol. in cholate. In the solubilized form, R.cntdot.MSH was extremely stable and dissocd. to an extent of only 30% within 12 days at 4.degree.. While this high stability can be maintained in the pH range of 7.0-8.5, the solubilized R.cntdot.MSH complex becomes increasingly unstable below pH 7.0 and totally dissocs. at a pH <6.0. In the membrane-bound form, the R.cntdot.MSH complex shows a parallel pH stability profile which is shifted down by approx. two pH units. In addn. to low pH, the R.cntdot.MSH complex becomes

unstable and totally dissocs. in the presence of 10 mM EGTA, suggesting that the calcium-sensitive function of the receptor is still assocd. with the receptor in the detergent-sol. state. The R.cntdot.MSH complexes in the sol. and membrane-bound forms are also totally resistant to proteolytic digestion by V8 protease, but were slowly digested by trypsin. Treatment of R.cntdot.MSH with 1-ethyl-3-(3-dimethylamino-propyl)carbodiimide hydrochloride or bis(sulfosuccinimidyl) suberate led to covalent crosslinking of MSH to the receptor mol. The electrophoretic mobility on SDS-PAGE of the 43/**46 kDa** doublet of the receptor-MSH conjugate (R*MSH) was identical to the photoaffinity labeled MSH receptor product described earlier in cultured M2R cells. However, the efficiency of prodn. of the crosslinked product was .apprx.30%, much higher than that achieved previously by photoaffinity labeling. Using rabbit polyclonal anti-.alpha.MSH antibodies, the R*MSH conjugate was identifiable on Western immunoblots. These results provide a basis for further development of procedures for purifn. of the MSH receptor mol. and studying its protein structure.

CC 2-1 (Mammalian Hormones)
Section cross-reference(s): 14

ST MSH receptor **melanoma** solubilization

IT Receptors
RL: PROC (Process)
(MSH, from **melanoma**, solubilization of, in complexes with ligand and crosslinked conjugate)

IT **Melanoma**
(amelanotic, MSH receptor of, solubilization of)

IT **Melanoma**
(melanotic, MSH receptor of, solubilization of)

IT 9002-79-3, Melanocyte-stimulating hormone
RL: ANST (Analytical study)
(receptor for, solubilization of, from **melanomas** in complexes with ligand and crosslinked conjugate)

L6 ANSWER 8 OF 8 HCAPLUS COPYRIGHT 1998 ACS

AN 1990:53087 HCAPLUS

DN 112:53087

TI Synthesis of vitamin K-dependent proteins by cultured human **tumor** cells

AU Al-Mondhiry, Hamid; Wallin, Reidar

CS Coll. Med., Pennsylvania State Univ., Hershey, PA, 17033, USA

SO Thromb. Haemostasis (1989), 62(2), 661-6
CODEN: THHADQ; ISSN: 0340-6245

DT Journal

LA English

AB The observation that warfarin inhibits the growth and metastasis of certain types of clin. and exptl. tumors suggests a role for vitamin K in tumor biol. The synthesis of vitamin K-dependent proteins was investigated in four malignant (lung epidermoid carcinoma, melanoma, colon adenocarcinoma, and breast adenocarcinoma) and three normal (colon epithelium, breast epithelium, and fibroblasts) cell lines of human origin grown in tissue cultures. The results show the following: 1) vitamin K-dependent carboxylase activity is present in all of the malignant and normal cell lines studied; 2) the malignant, as well as normal, cell lines synthesize a family of vitamin K-dependent proteins, and microsomal precursors of these proteins with apparent mol. masses of 74, 62, and **34**

kDa are common to all malignant and normal cell lines, whereas precursors of higher and lower mol. mass seem to be synthesized by some but not all tumor cell lines; and 3) the 74-kDa precursor synthesized by colon carcinoma and breast carcinoma was pos. identified as a precursor of protein S.

CC 14-1 (Mammalian Pathological Biochemistry)
Section cross-reference(s): 13

ST vitamin K protein formation **tumor** cell

IT **Melanoma**
Neoplasm, metabolism
(vitamin K-dependent protein formation by, of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)

IT Intestine, neoplasm
(colon, adenocarcinoma, vitamin K-dependent protein formation by, of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)

IT Lung, neoplasm
(epidermoid carcinoma, vitamin K-dependent protein formation by, of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)

IT Mammary gland
(neoplasm, adenocarcinoma, vitamin K-dependent protein formation by, of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)

IT Proteins, specific or class
RL: BIOL (Biological study)
(.gamma.-carboxyglutamic acid-contg., 34,000-mol.-wt., formation of vitamin K-dependent, by neoplastic and normal cells of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)

IT Proteins, specific or class
RL: BIOL (Biological study)
(.gamma.-carboxyglutamic acid-contg., 62,000-mol.-wt., formation of vitamin K-dependent, by neoplastic and normal cells of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)

IT Proteins, specific or class
RL: BIOL (Biological study)
(.gamma.-carboxyglutamic acid-contg., 74,000-mol.-wt., formation of vitamin K-dependent, by neoplastic and normal cells of humans, **tumor** growth and metastasis and **tumor**-assocd. hypercoagulable states in relation to)

****

****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:43:30 1998; MasPar time 2.22 Seconds
34.710 Million cell updates/sec
Tabular output not generated.

Title: >US-08-190-411A-2
Description: (1-14) from 5541104.pep
Perfect Score: 93
Sequence: 1 INFTRQRQPSEGSS 14

Scoring table: PAM 150
Gap 15

Searched: 60183 seqs, 5492030 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: a-issued
1:back1 2:51 3:52 4:53 5:54 6:55 7:56 8:57 9:PCT90
10:PCT91 11:PCT92 12:PCT93 13:PCT94 14:PCT95 15:PCT96

Statistics: Mean 16.616; Variance 48.328; scale 0.344

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	93	100.0	14	6	US-08-190-	Sequence 2, Applicatio	1.51e-04
2	53	57.0	420	7	US-08-391-	Sequence 10, Applicati	1.13e+01
3	53	57.0	420	7	US-08-391-	Sequence 2, Applicatio	1.13e+01
4	53	57.0	599	7	US-08-463-	Sequence 3, Applicatio	1.13e+01
5	53	57.0	613	7	US-08-405-	Sequence 1, Applicatio	1.13e+01
6	53	57.0	614	14	PCT-US95-0	Sequence 1, Applicatio	1.13e+01
7	53	57.0	614	7	US-08-225-	Sequence 1, Applicatio	1.13e+01
8	53	57.0	637	6	US-08-235-	Sequence 16, Applicati	1.13e+01
9	53	57.0	637	6	US-08-235-	Sequence 14, Applicati	1.13e+01
10	51	54.8	834	5	US-07-977-	Sequence 8, Applicatio	1.90e+01
11	51	54.8	834	10	PCT-US91-0	Sequence 8, Applicatio	1.90e+01
12	50	53.8	311	6	US-07-917-	Sequence 6, Applicatio	2.46e+01
13	50	53.8	311	7	US-08-479-	Sequence 6, Applicatio	2.46e+01
14	50	53.8	311	6	US-07-917-	Sequence 5, Applicatio	2.46e+01
15	50	53.8	311	7	US-08-479-	Sequence 5, Applicatio	2.46e+01
16	50	53.8	889	7	US-08-118-	Sequence 4, Applicatio	2.46e+01
17	48	51.6	355	7	US-08-196-	Sequence 34, Applicati	4.09e+01
18	48	51.6	355	8	US-08-681-	Sequence 34, Applicati	4.09e+01
19	47	50.5	369	6	US-07-688-	Sequence 32, Applicati	5.26e+01
20	47	50.5	369	10	PCT-US91-0	Sequence 31, Applicati	5.26e+01
21	47	50.5	527	14	PCT-US95-0	Sequence 10, Applicati	5.26e+01
22	47	50.5	599	6	US-08-391-	Sequence 4, Applicatio	5.26e+01
23	46	49.5	249	13	PCT-US94-0	Sequence 2, Applicatio	6.75e+01
24	46	49.5	303	7	US-08-479-	Sequence 3, Applicatio	6.75e+01
25	46	49.5	303	6	US-07-917-	Sequence 2, Applicatio	6.75e+01
26	46	49.5	303	6	US-07-917-	Sequence 3, Applicatio	6.75e+01
27	46	49.5	303	7	US-08-479-	Sequence 2, Applicatio	6.75e+01
28	46	49.5	830	5	US-07-977-	Sequence 6, Applicatio	6.75e+01
29	46	49.5	830	10	PCT-US91-0	Sequence 6, Applicatio	6.75e+01
30	46	49.5	832	5	US-07-977-	Sequence 2, Applicatio	6.75e+01
31	46	49.5	832	5	US-08-156-	Sequence 6, Applicatio	6.75e+01
32	46	49.5	832	5	US-08-156-	Sequence 8, Applicatio	6.75e+01
33	46	49.5	832	5	US-08-156-	Sequence 4, Applicatio	6.75e+01
34	46	49.5	832	6	US-08-073-	Sequence 4, Applicatio	6.75e+01
35	46	49.5	832	7	US-08-254-	Sequence 4, Applicatio	6.75e+01
36	46	49.5	832	5	US-08-156-	Sequence 10, Applicati	6.75e+01
37	46	49.5	832	5	US-08-156-	Sequence 2, Applicatio	6.75e+01
38	46	49.5	832	7	US-08-483-	Sequence 4, Applicatio	6.75e+01
39	46	49.5	832	10	PCT-US91-0	Sequence 2, Applicatio	6.75e+01
40	45	48.4	35	6	US-08-168-	Sequence 9, Applicatio	8.65e+01
41	45	48.4	551	1	5198359-2	Patent No. 5198359.	8.65e+01
42	45	48.4	551	1	5449756-2	Patent No. 5449756.	8.65e+01
43	45	48.4	880	8	US-08-445-	Sequence 10, Applicati	8.65e+01
44	44	47.3	16	13	PCT-US94-0	Sequence 7, Applicatio	1.11e+02
45	44	47.3	233	7	US-08-221-	Sequence 7, Applicatio	1.11e+02
46	44	47.3	318	7	US-08-379-	Sequence 9, Applicatio	1.11e+02
47	44	47.3	318	4	US-07-704-	Sequence 9, Applicatio	1.11e+02
48	44	47.3	364	7	US-08-318-	Sequence 6, Applicatio	1.11e+02
49	44	47.3	454	1	5260432-2	Patent No. 5260432.	1.11e+02
50	44	47.3	501	14	PCT-US95-0	Sequence 4, Applicatio	1.11e+02
51	44	47.3	501	7	US-08-331-	Sequence 4, Applicatio	1.11e+02
52	44	47.3	501	8	US-08-250-	Sequence 4, Applicatio	1.11e+02
53	44	47.3	602	1	5432081-2	Patent No. 5432081.	1.11e+02
54	44	47.3	652	7	US-08-318-	Sequence 8, Applicatio	1.11e+02
55	44	47.3	824	7	US-08-221-	Sequence 3, Applicatio	1.11e+02
56	44	47.3	831	6	US-08-073-	Sequence 5, Applicatio	1.11e+02
57	44	47.3	831	7	US-08-254-	Sequence 5, Applicatio	1.11e+02
58	44	47.3	831	7	US-08-483-	Sequence 5, Applicatio	1.11e+02
59	44	47.3	997	14	PCT-US95-0	Sequence 1, Applicatio	1.11e+02
60	44	47.3	997	8	US-08-428-	Sequence 1, Applicatio	1.11e+02
61	44	47.3	997	5	US-08-232-	Sequence 1, Applicatio	1.11e+02
62	44	47.3	998	6	US-08-233-	Sequence 6, Applicatio	1.11e+02
63	44	47.3	1021	6	US-08-233-	Sequence 2, Applicatio	1.11e+02
64	43	46.2	99	6	US-08-036-	Sequence 26, Applicati	1.41e+02
65	43	46.2	119	1	5242798-3	Patent No. 5242798.	1.41e+02
66	43	46.2	375	7	US-08-446-	Sequence 2, Applicatio	1.41e+02
67	43	46.2	375	6	US-08-442-	Sequence 2, Applicatio	1.41e+02
68	43	46.2	375	7	US-08-444-	Sequence 2, Applicatio	1.41e+02
69	43	46.2	410	6	US-08-123-	Sequence 7, Applicatio	1.41e+02
70	43	46.2	410	6	US-08-123-	Sequence 5, Applicatio	1.41e+02
71	43	46.2	420	7	US-08-391-	Sequence 7, Applicatio	1.41e+02
72	43	46.2	420	7	US-08-391-	Sequence 11, Applicati	1.41e+02
73	43	46.2	643	12	PCT-US93-1	Sequence 4, Applicatio	1.41e+02
74	43	46.2	644	7	US-08-206-	Sequence 2, Applicatio	1.41e+02
75	43	46.2	678	12	PCT-US93-0	Sequence 3, Applicatio	1.41e+02
76	43	46.2	1336	12	PCT-US93-1	Sequence 6, Applicatio	1.41e+02
77	43	46.2	1898	7	US-08-056-	Sequence 94, Applicati	1.41e+02
78	42	45.2	159	7	US-08-459-	Sequence 14, Applicati	1.80e+02
79	42	45.2	159	8	US-08-459-	Sequence 14, Applicati	1.80e+02
80	42	45.2	159	6	US-08-197-	Sequence 14, Applicati	1.80e+02
81	42	45.2	240	14	PCT-US95-0	Sequence 77, Applicati	1.80e+02
82	42	45.2	240	14	PCT-US95-0	Sequence 76, Applicati	1.80e+02
83	42	45.2	240	7	US-08-261-	Sequence 76, Applicati	1.80e+02
84	42	45.2	240	7	US-08-261-	Sequence 77, Applicati	1.80e+02
85	42	45.2	364	8	US-08-459-	Sequence 29, Applicati	1.80e+02
86	42	45.2	364	7	US-08-459-	Sequence 29, Applicati	1.80e+02
87	42	45.2	364	6	US-08-197-	Sequence 29, Applicati	1.80e+02
88	42	45.2	821	7	US-08-339-	Sequence 2, Applicatio	1.80e+02
89	42	45.2	833	7	US-08-483-	Sequence 8, Applicatio	1.80e+02
90	42	45.2	833	7	US-08-254-	Sequence 8, Applicatio	1.80e+02
91	42	45.2	833	6	US-08-073-	Sequence 8, Applicatio	1.80e+02
92	42	45.2	1873	8	US-08-435-	Sequence 4, Applicatio	1.80e+02
93	42	45.2	2101	12	PCT-US93-0	Sequence 4, Applicatio	1.80e+02
94	42	45.2	2101	7	US-08-470-	Sequence 4, Applicatio	1.80e+02
95	42	45.2	2101	7	US-08-466-	Sequence 4, Applicatio	1.80e+02

96	41	44.1	738	7	US-08-530-	Sequence 3, Applicatio	2.29e+02
97	41	44.1	738	7	US-08-530-	Sequence 11, Applicati	2.29e+02
98	41	44.1	869	7	US-08-646-	Sequence 32, Applicati	2.29e+02
99	41	44.1	890	8	US-08-445-	Sequence 2, Applicatio	2.29e+02
100	41	44.1	1203	12	PCT-US93-1	Sequence 103, Applicat	2.29e+02

ALIGNMENTS

```
RESULT    1
ID   US-08-190-411A-2     STANDARD;     PRT;   14 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/08190411A.
XX
CC    Sequence 2, Application US/08190411A
CC    Patent No. 5541104
CC     GENERAL INFORMATION:
CC       APPLICANT:  Chen, Yao-Tseng; Stockert, Elisabeth;
CC       APPLICANT:  Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.;
CC       APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry;
CC       APPLICANT:  Old, Lloyd J.
CC       TITLE OF INVENTION:  MONOCLONAL ANTIBODIES WHICH BIND TO
CC       TITLE OF INVENTION:  TUMOR REJECTION ANTIGEN PRECURSOR MAGE-1, RECOMBI
  NANT MAGE-1,
CC       TITLE OF INVENTION:  AND MAGE-1 DERIVED IMMUNOGENIC PEPTIDES
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/190,411A
CC         FILING DATE:  01-FEBRUARY-1994
CC         CLASSIFICATION:  436
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  037,230
CC         FILING DATE:  26-MARCH-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/04354
CC         FILING DATE:  22-MAY-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/807,043
CC         FILING DATE:  12-DECEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/764,364
CC         FILING DATE:  23-SEPTEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/728,838
CC         APPLICATION NUMBER:  9-JULY-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/705,702
CC         FILING DATE:  23-MAY-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5541104man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5354
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  14 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   14 AA; 1607 MW; 1134 CN;

  Query Match          100.0%;  Score 93;  DB 6;  Length 14;
  Best Local Similarity 100.0%;  Pred. No. 1.51e-04;
  Matches   14;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 INFTRQRQPSEGSS 14
          ||||||||||||||
Qy      1 INFTRQRQPSEGSS 14


RESULT    2
ID   US-08-391-259-10     STANDARD;     PRT;   420 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 10, Application US/08391259.
XX
CC    Sequence 10, Application US/08391259
CC    Patent No. 5621078
CC     GENERAL INFORMATION:
CC       APPLICANT:  Riemen, Mark W
CC       APPLICANT:  Stirdivant, Steven M
CC       TITLE OF INVENTION:  Modified PE40
CC       NUMBER OF SEQUENCES:  11
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Merck & Co., Inc.
CC         STREET:  126 Lincoln Avenue
CC         CITY:  Rahway
CC         STATE:  New Jersey
CC         COUNTRY:  U.S.
CC         ZIP:  07065
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/391,259
CC         FILING DATE:
CC         CLASSIFICATION:  530
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/120,698
CC         FILING DATE:
CC         APPLICATION NUMBER:  US/07/879,037
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Grassler, Frank P
CC         REGISTRATION NUMBER:  31,164
CC         REFERENCE/DOCKET NUMBER:  17879IA
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (908) 594-3462
CC         TELEFAX:  (908) 594-4720
CC     INFORMATION FOR SEQ ID NO:  10:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  420 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   420 AA; 45129 MW; 823636 CN;

  Query Match          57.0%;  Score 53;  DB 7;  Length 420;
  Best Local Similarity 85.7%;  Pred. No. 1.13e+01;
  Matches    6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db     79 FTRHRQP 85
          |||:|||
```

```
Qy      3 FTRQRQP 9


RESULT    3
ID   US-08-391-259-2     STANDARD;     PRT;   420 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/08391259.
XX
CC    Sequence 2, Application US/08391259
CC    Patent No. 5621078
CC     GENERAL INFORMATION:
CC       APPLICANT:  Riemen, Mark W
CC       APPLICANT:  Stirdivant, Steven M
CC       TITLE OF INVENTION:  Modified PE40
CC       NUMBER OF SEQUENCES:  11
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Merck & Co., Inc.
CC         STREET:  126 Lincoln Avenue
CC         CITY:  Rahway
CC         STATE:  New Jersey
CC         COUNTRY:  U.S.
CC         ZIP:  07065
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/391,259
CC         FILING DATE:
CC         CLASSIFICATION:  530
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/120,698
CC         FILING DATE:
CC         APPLICATION NUMBER:  US/07/879,037
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Grassler, Frank P
CC         REGISTRATION NUMBER:  31,164
CC         REFERENCE/DOCKET NUMBER:  17879IA
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (908) 594-3462
CC         TELEFAX:  (908) 594-4720
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  420 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   420 AA; 45179 MW; 824920 CN;

  Query Match           57.0%;  Score 53;  DB 7;  Length 420;
  Best Local Similarity  85.7%;  Pred. No. 1.13e+01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db     79 FTRHRQP 85
          |||:|||
Qy      3 FTRQRQP 9


RESULT    4
ID   US-08-463-163-3     STANDARD;     PRT;   599 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 3, Application US/08463163.
XX
CC    Sequence 3, Application US/08463163
CC    Patent No. 5696237
CC     GENERAL INFORMATION:
CC       APPLICANT:  FitzGerald, David J.
CC       APPLICANT:  Chaudhary, Vijay K.
CC       APPLICANT:  Pastan, Ira H.
CC       APPLICANT:  Waldmann, Thomas A.
CC       APPLICANT:  Queen, Cary L.
CC       TITLE OF INVENTION:  Recombinant Antibody-Toxin Fusion Protein
CC       NUMBER OF SEQUENCES:  12
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Townsend and Townsend and Crew
CC         STREET:  One Market Plaza, Steuart Street Tower
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  94105-1492
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/463,163
CC         FILING DATE:  05-JUN-1995
CC         CLASSIFICATION:  536
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 06/227,227
CC         FILING DATE:  22-JAN-1981
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 06/911,227
CC         FILING DATE:  24-SEP-1986
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/341,361
CC         FILING DATE:  21-APR-1989
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/865,722
CC         FILING DATE:  08-APR-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Ellen L.
CC         REGISTRATION NUMBER:  32,762
CC         REFERENCE/DOCKET NUMBER:  015280-12211
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 543-9600
CC         TELEFAX:  (415) 543-5043
CC     INFORMATION FOR SEQ ID NO:  3:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  599 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   599 AA; 64018 MW; 1732802 CN;

  Query Match           57.0%;  Score 53;  DB 7;  Length 599;
  Best Local Similarity  85.7%;  Pred. No. 1.13e+01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    258 FTRHRQP 264
          |||:|||
Qy      3 FTRQRQP 9


RESULT    5
ID   US-08-405-615-1     STANDARD;     PRT;   613 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application US/08405615.
XX
CC    Sequence 1, Application US/08405615
```

```
CC    Patent No. 5602095
CC     GENERAL INFORMATION:
CC       APPLICANT:  Pastan, Ira
CC       APPLICANT:  FitzGerald, David J.
CC       TITLE OF INVENTION:  Recombinant Pseudomonas Exotoxin with
CC       TITLE OF INVENTION:  Increased Activity
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Ellen L. Weber
CC         STREET:  One Market Plaza, Steuart Tower, Suite 2000
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/405,615
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/901,709
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Ellen L.
CC         REGISTRATION NUMBER:  32,762
CC         REFERENCE/DOCKET NUMBER:  15280-36
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-543-9600
CC         TELEFAX:  415-543-5043
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  613 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   613 AA; 66827 MW; 1767040 CN;

  Query Match           57.0%;  Score 53;  DB 7;  Length 613;
  Best Local Similarity  85.7%;  Pred. No. 1.13e+01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    272 FTRHRQP 278
          |||:|||
Qy      3 FTRQRQP 9


RESULT    6
ID   PCT-US95-04468-1     STANDARD;     PRT;   614 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application PC/TUS9504468.
XX
CC    Sequence 1, Application PC/TUS9504468
CC     GENERAL INFORMATION:
CC       APPLICANT:
CC       TITLE OF INVENTION:  CIRCULARLY PERMUTATED LIGANDS AND
CC       TITLE OF INVENTION:  CIRCULARLY PERMUTED FUSION PROTEINS
CC       NUMBER OF SEQUENCES:  59
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US95/04468
CC         FILING DATE:  07-APR-1995
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/225,224
CC         FILING DATE:  08-APR-1994
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Ellen L.
CC         REGISTRATION NUMBER:  32,762
CC         REFERENCE/DOCKET NUMBER:  15280-193-1PC
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 543-9600
CC         TELEFAX:  (415) 543-5043
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  614 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  unknown
CC         TOPOLOGY:  unknown
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  Protein
CC         LOCATION:  1..614
CC         OTHER INFORMATION:  /label= native-PE
SQ   SEQUENCE   614 AA; 66959 MW; 1772843 CN;

  Query Match           57.0%;  Score 53;  DB 14;  Length 614;
  Best Local Similarity  85.7%;  Pred. No. 1.13e+01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    273 FTRHRQP 279
          |||:|||
Qy      3 FTRQRQP 9


RESULT    7
ID   US-08-225-224-1     STANDARD;     PRT;   614 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application US/08225224.
XX
CC    Sequence 1, Application US/08225224
CC    Patent No. 5635599
CC     GENERAL INFORMATION:
CC       APPLICANT:  PASTAN, Ira
CC       APPLICANT:  KREITMAN, Robert J.
CC       TITLE OF INVENTION:  CIRCULARLY PERMUTATED LIGANDS AND
CC       TITLE OF INVENTION:  CIRCULARLY PERMUTED FUSION PROTEINS
CC       NUMBER OF SEQUENCES:  57
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Townsend and Townsend Khourie and Crew
CC         STREET:  Steuart Street Tower, One Market Plaza
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  US
CC         ZIP:  94105-1493
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/225,224
CC         FILING DATE:  8-APR-1994
CC         CLASSIFICATION:  530
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Ellen L.
CC         REGISTRATION NUMBER:  32,762
CC         REFERENCE/DOCKET NUMBER:  15280-193
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 543-9600
```

```
CC         TELEFAX:  (415) 543-5043
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  614 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  unknown
CC         TOPOLOGY:  unknown
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  Protein
CC         LOCATION:  1..614
CC         OTHER INFORMATION:  /label= native-PE
SQ   SEQUENCE   614 AA; 66959 MW; 1772843 CN;

  Query Match           57.0%;  Score 53;  DB 7;  Length 614;
  Best Local Similarity  85.7%;  Pred. No. 1.13e+01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    273 FTRHRQP 279
          |||:|||
Qy      3 FTRQRQP 9


RESULT    8
ID   US-08-235-838-16     STANDARD;     PRT;   637 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 16, Application US/08235838.
XX
CC    Sequence 16, Application US/08235838
CC    Patent No. 5571894
CC     GENERAL INFORMATION:
CC       APPLICANT:  Wels, Winfried S.
CC       APPLICANT:  Hynes, Nancy E.
CC       APPLICANT:  Harwerth, Ina-Maria
CC       APPLICANT:  Groner, Bernd
CC       APPLICANT:  Hardman, No. 5571894man
CC       APPLICANT:  Zwickl, Markus
CC       TITLE OF INVENTION:  Recombinant Antibodies Specific for a
CC       TITLE OF INVENTION:  Growth Factor Receptor
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  CIBA-GEIGY Corporation
CC         STREET:  7 Skyline Drive
CC         CITY:  Hawthorne
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10532
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/235,838
CC         FILING DATE:  TBA
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/828,832
CC         FILING DATE:  31-JAN-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  GB 91-810079.3
CC         FILING DATE:  05-FEB-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Elmer, James Scott
CC         REGISTRATION NUMBER:  36,129
CC         REFERENCE/DOCKET NUMBER:  4-18518/A/CIP/CONT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (919)541-8614
CC         TELEFAX:  (919)541-8689
CC     INFORMATION FOR SEQ ID NO:  16:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  637 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   637 AA; 68361 MW; 1955217 CN;

  Query Match           57.0%;  Score 53;  DB 6;  Length 637;
  Best Local Similarity  85.7%;  Pred. No. 1.13e+01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    296 FTRHRQP 302
          |||:|||
Qy      3 FTRQRQP 9


RESULT    9
ID   US-08-235-838-14     STANDARD;     PRT;   637 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 14, Application US/08235838.
XX
CC    Sequence 14, Application US/08235838
CC    Patent No. 5571894
CC     GENERAL INFORMATION:
CC       APPLICANT:  Wels, Winfried S.
CC       APPLICANT:  Hynes, Nancy E.
CC       APPLICANT:  Harwerth, Ina-Maria
CC       APPLICANT:  Groner, Bernd
CC       APPLICANT:  Hardman, No. 5571894man
CC       APPLICANT:  Zwickl, Markus
CC       TITLE OF INVENTION:  Recombinant Antibodies Specific for a
CC       TITLE OF INVENTION:  Growth Factor Receptor
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  CIBA-GEIGY Corporation
CC         STREET:  7 Skyline Drive
CC         CITY:  Hawthorne
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10532
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/235,838
CC         FILING DATE:  TBA
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/828,832
CC         FILING DATE:  31-JAN-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  GB 91-810079.3
CC         FILING DATE:  05-FEB-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Elmer, James Scott
CC         REGISTRATION NUMBER:  36,129
CC         REFERENCE/DOCKET NUMBER:  4-18518/A/CIP/CONT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (919)541-8614
CC         TELEFAX:  (919)541-8689
CC     INFORMATION FOR SEQ ID NO:  14:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  637 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
```

```
SQ   SEQUENCE   637 AA; 68441 MW; 1957762 CN;

  Query Match           57.0%;  Score 53;  DB 6;  Length 637;
  Best Local Similarity  85.7%;  Pred. No. 1.13e+01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    296 FTRHRQP 302
          |||:|||
Qy      3 FTRQRQP 9


RESULT   10
ID   US-07-977-434-8     STANDARD;     PRT;   834 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/07977434.
XX
CC    Sequence 8, Application US/07977434
CC    Patent No. 5466591
CC     GENERAL INFORMATION:
CC       APPLICANT:  Gelfand, David H.
CC       APPLICANT:  Abramson, Richard D.
CC       TITLE OF INVENTION:  5' TO 3' EXONUCLEASE MUTATIONS OF
CC       TITLE OF INVENTION:  THERMOSTABLE DNA POLYMERASES
CC       NUMBER OF SEQUENCES:  38
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Hoffmann-La Roche Inc.
CC         STREET:  340 Kingsland Street
CC         CITY:  Nutley
CC         STATE:  New Jersey
CC         ZIP:  07110-1199
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  Macintosh
CC         OPERATING SYSTEM:  7
CC         SOFTWARE:  WordPerfect 2.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/977,434
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,490
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,466
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,213
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 523,394
CC         FILING DATE:  15-MAY-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 143,441
CC         FILING DATE:  12-JAN-1988
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 063,509
CC         FILING DATE:  17-JUN-1987
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 899,241
CC         FILING DATE:  22-AUG-1986
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 746,121
CC         FILING DATE:  15-AUG-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  WO PCT/US90/07641
CC         FILING DATE:  21-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 585,471
CC         FILING DATE:  20-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 455,611
CC         FILING DATE:  22-DEC-1989
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 609,157
CC         FILING DATE:  02-NOV-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 557,517
CC         FILING DATE:  24-JUL-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Luann Cserr
CC         REGISTRATION NUMBER:  31,822
CC         REFERENCE/DOCKET NUMBER:  Case No. 5466591 8753
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (510) 814-2972
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  834 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   834 AA; 94055 MW; 3318124 CN;

  Query Match           54.8%;  Score 51;  DB 5;  Length 834;
  Best Local Similarity  44.4%;  Pred. No. 1.90e+01;
  Matches     4;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db    258 VDFARRREP 266
          ::|:|:|:|
Qy      1 INFTRQRQP 9


RESULT   11
ID   PCT-US91-07035-8     STANDARD;     PRT;   834 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application PC/TUS9107035.
XX
CC    Sequence 8, Application PC/TUS9107035
CC     GENERAL INFORMATION:
CC       APPLICANT:  Gelfand, David H.
CC       APPLICANT:  Abramson, Richard D.
CC       TITLE OF INVENTION:  5' TO 3' EXONUCLEASE MUTATIONS OF
CC       TITLE OF INVENTION:  THERMOSTABLE DNA POLYMERASES
CC       NUMBER OF SEQUENCES:  38
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cetus Corporation
CC         STREET:  1400 Fifty-third Street
CC         CITY:  Emeryville
CC         STATE:  California
CC         ZIP:  94608
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  WordPerfect 5.0
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/07035
CC         FILING DATE:  19910930
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,490
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,466
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,213
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
```

```
CC         APPLICATION NUMBER:  US 523,394
CC         FILING DATE:  15-MAY-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 143,441
CC         FILING DATE:  12-JAN-1988
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 063,509
CC         FILING DATE:  17-JUN-1987
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 899,241
CC         FILING DATE:  22-AUG-1986
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 746,121
CC         FILING DATE:  15-AUG-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  WO PCT/US90/07641
CC         FILING DATE:  21-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 585,471
CC         FILING DATE:  20-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 455,611
CC         FILING DATE:  22-DEC-1989
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 609,157
CC         FILING DATE:  02-NOV-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 557,517
CC         FILING DATE:  24-JUL-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Sias Ph.D, Stacey R.
CC         REGISTRATION NUMBER:  32,630
CC         REFERENCE/DOCKET NUMBER:  Case No. 2580
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-420-3300
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  834 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   834 AA; 94055 MW; 3318124 CN;

  Query Match           54.8%;  Score 51;  DB 10;  Length 834;
  Best Local Similarity  44.4%;  Pred. No. 1.90e+01;
  Matches    4;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db    258 VDFARRREP 266
          ::|:|:|:|
Qy      1 INFTRQRQP 9


RESULT   12
ID   US-07-917-111-6     STANDARD;     PRT;   311 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application US/07917111.
XX
CC    Sequence 6, Application US/07917111
CC    Patent No. 5565344
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nanba, Hirokazu
CC       APPLICANT:  Yamada, Yukio
CC       APPLICANT:  Takano, Masayuki
CC       APPLICANT:  Ikenaka, Yasuhiro
CC       APPLICANT:  Takahashi, Satomi
CC       APPLICANT:  Yajima, Kazuyoshi
CC       TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF D-ALPHA-AMINO
CC       TITLE OF INVENTION:  ACIDS
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Wegner, Cantor, Mueller & Player
CC         STREET:  1233 20th Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         ZIP:  20036-8218
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/917,111
CC         FILING DATE:  19920807
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 400848/1990
CC         FILING DATE:  07-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 407922/1990
CC         FILING DATE:  27-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 078840/1991
CC         FILING DATE:  11-APR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/JP91/01696
CC         FILING DATE:  06-DEC-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Player Esq., William E.
CC         REGISTRATION NUMBER:  31,409
CC         REFERENCE/DOCKET NUMBER:  P-500-23486
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-887-0400
CC         TELEFAX:  202-835-0605
CC         TELEX:  440706
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  311 amino acids
CC         TYPE:  AMINO ACID
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Pseudomonas
CC         STRAIN:  KNK 003A (FERM BP-3181)
SQ   SEQUENCE   311 AA; 35307 MW; 480721 CN;

  Query Match           53.8%;  Score 50;  DB 6;  Length 311;
  Best Local Similarity  44.4%;  Pred. No. 2.46e+01;
  Matches    4;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db    286 FDFARHREP 294
          ::|:|:|:|
Qy      1 INFTRQRQP 9


RESULT   13
ID   US-08-479-638-6     STANDARD;     PRT;   311 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application US/08479638.
XX
CC    Sequence 6, Application US/08479638
CC    Patent No. 5695968
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nanba, Hirokazu
CC       APPLICANT:  Yamada, Yukio
CC       APPLICANT:  Takano, Masayuki
CC       APPLICANT:  Ikenaka, Yasuhiro
CC       APPLICANT:  Takahashi, Satomi
```

```
CC       APPLICANT:  Yajima, Kazuyoshi
CC       TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF D-ALPHA-AMINO
CC       TITLE OF INVENTION:  ACIDS
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Wegner, Cantor, Mueller & Player
CC         STREET:  1233 20th Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         ZIP:  20036-8218
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/479,638
CC         FILING DATE:  07-JUN-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/917,111
CC         FILING DATE:  07-AUG-1992
CC         APPLICATION NUMBER:  JP 400848/1990
CC         FILING DATE:  07-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 407922/1990
CC         FILING DATE:  27-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 078840/1991
CC         FILING DATE:  11-APR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/JP91/01696
CC         FILING DATE:  06-DEC-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Player Esq., William E.
CC         REGISTRATION NUMBER:  31,409
CC         REFERENCE/DOCKET NUMBER:  P-500-23486
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-887-0400
CC         TELEFAX:  202-835-0605
CC         TELEX:  440706
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  311 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Pseudomonas
CC         STRAIN:  KNK 003A (FERM BP-3181)
SQ   SEQUENCE   311 AA; 35307 MW; 480721 CN;

  Query Match           53.8%;  Score 50;  DB 7;  Length 311;
  Best Local Similarity  44.4%;  Pred. No. 2.46e+01;
  Matches    4;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db    286 FDFARHREP 294
          ::|:|:|:|
Qy      1 INFTRQRQP 9


RESULT   14
ID   US-07-917-111-5     STANDARD;     PRT;   311 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application US/07917111.
XX
CC    Sequence 5, Application US/07917111
CC    Patent No. 5565344
```

```
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nanba, Hirokazu
CC       APPLICANT:  Yamada, Yukio
CC       APPLICANT:  Takano, Masayuki
CC       APPLICANT:  Ikenaka, Yasuhiro
CC       APPLICANT:  Takahashi, Satomi
CC       APPLICANT:  Yajima, Kazuyoshi
CC       TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF D-ALPHA-AMINO
CC       TITLE OF INVENTION:  ACIDS
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Wegner, Cantor, Mueller & Player
CC         STREET:  1233 20th Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         ZIP:  20036-8218
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/917,111
CC         FILING DATE:  19920807
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 400848/1990
CC         FILING DATE:  07-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 407922/1990
CC         FILING DATE:  27-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 078840/1991
CC         FILING DATE:  11-APR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/JP91/01696
CC         FILING DATE:  06-DEC-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Player Esq., William E.
CC         REGISTRATION NUMBER:  31,409
CC         REFERENCE/DOCKET NUMBER:  P-500-23486
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-887-0400
CC         TELEFAX:  202-835-0605
CC         TELEX:  440706
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  311 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   311 AA; 35307 MW; 480721 CN;

  Query Match           53.8%;  Score 50;  DB 6;  Length 311;
  Best Local Similarity  44.4%;  Pred. No. 2.46e+01;
  Matches    4;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db    286 FDFARHREP 294
          ::|:|:|:|
Qy      1 INFTRQRQP 9


RESULT   15
ID   US-08-479-638-5     STANDARD;     PRT;   311 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application US/08479638.
XX
CC    Sequence 5, Application US/08479638
CC    Patent No. 5695968
```

```
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nanba, Hirokazu
CC       APPLICANT:  Yamada, Yukio
CC       APPLICANT:  Takano, Masayuki
CC       APPLICANT:  Ikenaka, Yasuhiro
CC       APPLICANT:  Takahashi, Satomi
CC       APPLICANT:  Yajima, Kazuyoshi
CC       TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF D-ALPHA-AMINO
CC       TITLE OF INVENTION:  ACIDS
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Wegner, Cantor, Mueller & Player
CC         STREET:  1233 20th Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         ZIP:  20036-8218
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/479,638
CC         FILING DATE:  07-JUN-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/917,111
CC         FILING DATE:  07-AUG-1992
CC         APPLICATION NUMBER:  JP 400848/1990
CC         FILING DATE:  07-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 407922/1990
CC         FILING DATE:  27-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 078840/1991
CC         FILING DATE:  11-APR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/JP91/01696
CC         FILING DATE:  06-DEC-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Player Esq., William E.
CC         REGISTRATION NUMBER:  31,409
CC         REFERENCE/DOCKET NUMBER:  P-500-23486
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-887-0400
CC         TELEFAX:  202-835-0605
CC         TELEX:  440706
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  311 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   311 AA; 35307 MW; 480721 CN;

  Query Match           53.8%;  Score 50;  DB 7;  Length 311;
  Best Local Similarity  44.4%;  Pred. No. 2.46e+01;
  Matches     4;  Conservative     5;  Mismatches    0;  Indels    0;  Gaps    0;

Db    286 FDFARHREP 294
          ::|:|:|:|
Qy      1 INFTRQRQP 9


RESULT   16
ID   US-08-118-101A-4     STANDARD;     PRT;   889 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/08118101A.
XX
CC    Sequence 4, Application US/08118101A
CC    Patent No. 5620892
CC     GENERAL INFORMATION:
CC       APPLICANT:  Kurtz, Stephen E.
CC       APPLICANT:  Knickerbocker, Aron M.
CC       APPLICANT:  McCullough, John R.
CC       TITLE OF INVENTION:  A STRAIN OF SACCHAROMYCES CEREVISIAE
CC       TITLE OF INVENTION:  EXPRESSING THE GENE ENCODING POTASSIUM TRANSPORTE
  R MINK
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Burton Rodney
CC         STREET:  P.O. Box 4000
CC         CITY:  Princeton
CC         STATE:  New Jersey
CC         COUNTRY:  U.S.A.
CC         ZIP:  08543-4000
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/118,101A
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Gaul, Timothy J.
CC         REGISTRATION NUMBER:  33,111
CC         REFERENCE/DOCKET NUMBER:  DC27
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (609) 252-5901
CC         TELEFAX:  (609) 252-4526
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  889 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   889 AA; 101086 MW; 4294395 CN;

  Query Match           53.8%;  Score 50;  DB 7;  Length 889;
  Best Local Similarity  60.0%;  Pred. No. 2.46e+01;
  Matches     6;  Conservative     2;  Mismatches    2;  Indels    0;  Gaps    0;

Db    233 NFSSKRQPSD 242
          ||:  ||||:
Qy      2 NFTRQRQPSE 11


RESULT   17
ID   US-08-196-218-34     STANDARD;     PRT;   355 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 34, Application US/08196218.
XX
CC    Sequence 34, Application US/08196218
CC    Patent No. 5614619
CC     GENERAL INFORMATION:
CC       APPLICANT:  Piepersberg, Wolfgang
CC       APPLICANT:  Stockmann, Michael
CC       APPLICANT:  Taleghani, Kampiz Mansouri
CC       APPLICANT:  Distler, Jurgen
CC       APPLICANT:  Grabley, Susanne
CC       APPLICANT:  Sichel, Petra
CC       APPLICANT:  Brau, Barbara
CC       TITLE OF INVENTION:  Secondary-Metabolite Biosynthesis Genes
CC       TITLE OF INVENTION:  From Actinomycetes, Method of Isolating Them, and
   Their
CC       TITLE OF INVENTION:  Use.
```

```
CC       NUMBER OF SEQUENCES:  34
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
CC         ADDRESSEE:  Dunner
CC         STREET:  1300 I Street, N.W.
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  United States
CC         ZIP:  20005-3315
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/196,218
CC         FILING DATE:  25-AUG-1994
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Ogden, Stasia L.
CC         REGISTRATION NUMBER:  36,228
CC         REFERENCE/DOCKET NUMBER:  02481.1372-00000
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-408-4000
CC         TELEFAX:  202-408-4400
CC     INFORMATION FOR SEQ ID NO:  34:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  355 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   355 AA; 37616 MW; 642264 CN;

  Query Match           51.6%;  Score 48;  DB 7;  Length 355;
  Best Local Similarity  50.0%;  Pred. No. 4.09e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    120 EFTRQRPAAQ 129
          :|||||  :::
Qy      2 NFTRQRQPSE 11


RESULT   18
ID   US-08-681-953-34     STANDARD;     PRT;   355 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 34, Application US/08681953.
XX
CC    Sequence 34, Application US/08681953
CC    Patent No. 5710032
CC     GENERAL INFORMATION:
CC       APPLICANT:  Piepersberg, Wolfgang
CC       APPLICANT:  Stockmann, Michael
CC       APPLICANT:  Taleghani, Kampiz Mansouri
CC       APPLICANT:  Distler, Jurgen
CC       APPLICANT:  Grabley, Susanne
CC       APPLICANT:  Sichel, Petra
CC       APPLICANT:  Brau, Barbara
CC       TITLE OF INVENTION:  Secondary-Metabolite Biosynthesis Genes
CC       TITLE OF INVENTION:  From Actinomycetes, Method of Isolating Them, and
   Their
CC       TITLE OF INVENTION:  Use.
CC       NUMBER OF SEQUENCES:  34
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
CC         ADDRESSEE:  Dunner
CC         STREET:  1300 I Street, N.W.
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  United States
CC         ZIP:  20005-3315
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/681,953
CC         FILING DATE:  30-JUL-1996
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/196,218
CC         FILING DATE:  25-AUG-1994
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Ogden, Stasia L.
CC         REGISTRATION NUMBER:  36,228
CC         REFERENCE/DOCKET NUMBER:  02481.1372-00000
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-408-4000
CC         TELEFAX:  202-408-4400
CC     INFORMATION FOR SEQ ID NO:  34:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  355 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   355 AA; 37616 MW; 642264 CN;

  Query Match           51.6%;  Score 48;  DB 8;  Length 355;
  Best Local Similarity  50.0%;  Pred. No. 4.09e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    120 EFTRQRPAAQ 129
          :|||||  :::
Qy      2 NFTRQRQPSE 11


RESULT   19
ID   US-07-688-352C-32     STANDARD;     PRT;   369 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 32, Application US/07688352C.
XX
CC    Sequence 32, Application US/07688352C
CC    Patent No. 5527896
CC     GENERAL INFORMATION:
CC       APPLICANT:  Wigler, Michael H.
CC       APPLICANT:  Colicelli, John J.
CC       TITLE OF INVENTION:  Cloning by Complementation and Related
CC       TITLE OF INVENTION:  Processes
CC       NUMBER OF SEQUENCES:  57
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
CC         ADDRESSEE:  Bicknell
CC         STREET:  Two First National Plaza, 20 South Clark
CC         STREET:  Street
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/688,352C
CC         FILING DATE:  19910419
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
```

```
CC         APPLICATION NUMBER:  US 07/511,715
CC         FILING DATE:  20-APR-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Borun, Michael F.
CC         REGISTRATION NUMBER:  25447
CC         REFERENCE/DOCKET NUMBER:  27805/30197
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (312) 346-5750
CC         TELEFAX:  (312) 984-9740
CC         TELEX:  25-3856
CC     INFORMATION FOR SEQ ID NO:  32:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  369 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   369 AA; 40466 MW; 592164 CN;

  Query Match           50.5%;  Score 47;  DB 6;  Length 369;
  Best Local Similarity  60.0%;  Pred. No. 5.26e+01;
  Matches     6;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    296 RQRQPREDGN 305
          ||||| | ::
Qy      5 RQRQPSEGSS 14


RESULT   20
ID   PCT-US91-02714-31     STANDARD;     PRT;   369 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 31, Application PC/TUS9102714.
XX
CC    Sequence 31, Application PC/TUS9102714
CC     GENERAL INFORMATION:
CC       APPLICANT:  Wigler, Michael H.
CC       APPLICANT:  Colicelli, John J.
CC       TITLE OF INVENTION:  Cloning by Complementation and Related
CC       TITLE OF INVENTION:  Processes
CC       NUMBER OF SEQUENCES:  55
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
CC         ADDRESSEE:  Bicknell
CC         STREET:  Two First National Plaza, 20 South Clark
CC         STREET:  Street
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  USA
CC         ZIP:  60603
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/02714
CC         FILING DATE:  19910419
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/511,715
CC         FILING DATE:  20-APR-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Borun, Michael F.
CC         REGISTRATION NUMBER:  25447
CC         REFERENCE/DOCKET NUMBER:  27805/30197
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (312) 346-5750
CC         TELEFAX:  (312) 984-9740
CC         TELEX:  25-3856
CC     INFORMATION FOR SEQ ID NO:  31:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  369 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   369 AA; 40466 MW; 592164 CN;

  Query Match           50.5%;  Score 47;  DB 10;  Length 369;
  Best Local Similarity  60.0%;  Pred. No. 5.26e+01;
  Matches     6;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    296 RQRQPREDGN 305
          ||||| | ::
Qy      5 RQRQPSEGSS 14


RESULT   21
ID   PCT-US95-05008-10     STANDARD;     PRT;   527 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 10, Application PC/TUS9505008.
XX
CC    Sequence 10, Application PC/TUS9505008
CC     GENERAL INFORMATION:
CC       APPLICANT:  Sugen, Inc.
CC       APPLICANT:  515 Galveston Drive
CC       APPLICANT:  Redwood City, California 94063-4720
CC       APPLICANT:  United States of America
CC       APPLICANT:  Wissenschaften E.V.
CC       APPLICANT:  Hofgarten Str. 2
CC       APPLICANT:  Munchen 80539
CC       APPLICANT:  Germany
CC       TITLE OF INVENTION:  Novel Megakaryocytic Protein Tyrosine
CC       TITLE OF INVENTION:  Kinases
CC       NUMBER OF SEQUENCES:  21
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pennie & Edmonds
CC         STREET:  1155 Avenue of the Americas
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  U.S.A.
CC         ZIP:  10036
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US95/05008
CC         FILING DATE:  24-APR-1995
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/232,545
CC         FILING DATE:  22-APR-1994
CC         CLASSIFICATION:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Coruzzi, Laura A.
CC         REGISTRATION NUMBER:  30,742
CC         REFERENCE/DOCKET NUMBER:  7683-074
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212)790-9090
CC         TELEFAX:  (212)869-9741
CC         TELEX:  66141 PENNIE
CC     INFORMATION FOR SEQ ID NO:  10:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  527 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  unknown
CC         TOPOLOGY:  unknown
CC       MOLECULE TYPE:  protein
```

```
SQ   SEQUENCE   527 AA; 61556 MW; 1469994 CN;

  Query Match           50.5%;  Score 47;  DB 14;  Length 527;
  Best Local Similarity  75.0%;  Pred. No. 5.26e+01;
  Matches     6;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    347 LNFLRQRQ 354
          :|| ||||
Qy      1 INFTRQRQ 8


RESULT   22
ID   US-08-391-615-4     STANDARD;     PRT;   599 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/08391615.
XX
CC    Sequence 4, Application US/08391615
CC    Patent No. 5550054
CC     GENERAL INFORMATION:
CC       APPLICANT:  Witte, Owen
CC       APPLICANT:  Tsukada, Satoshi
CC       APPLICANT:  Saffran, Douglas
CC       APPLICANT:  Rawlings, David
CC       TITLE OF INVENTION:  HEMATOPOIETIC RESTRICTED TYROSINE KINASE
CC       TITLE OF INVENTION:
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
CC         STREET:  4 Embarcadero Center, Suite 3400
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  94111-4187
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/391,615
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/006,449
CC         FILING DATE:  21-JAN-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Rowland, Bertram I
CC         REGISTRATION NUMBER:  20,015
CC         REFERENCE/DOCKET NUMBER:  A-57583-1/BIR  UCLA 232-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 781-1989
CC         TELEFAX:  (415) 398-3249
CC         TELEX:  910 277299 FHT UR
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  599 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   599 AA; 69815 MW; 1897283 CN;

  Query Match           50.5%;  Score 47;  DB 6;  Length 599;
  Best Local Similarity  75.0%;  Pred. No. 5.26e+01;
  Matches     6;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    425 LNFLRQRQ 432
          :|| ||||
Qy      1 INFTRQRQ 8


RESULT   23
ID   PCT-US94-06430-2     STANDARD;     PRT;   249 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9406430.
XX
CC    Sequence 2, Application PC/TUS9406430
CC     GENERAL INFORMATION:
CC       APPLICANT:  The Upjohn Company
CC       TITLE OF INVENTION:  Lettuce Infectious Yellows Virus Genes
CC       NUMBER OF SEQUENCES:  25
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  The Upjohn Company, Corp. Intellectual
CC         ADDRESSEE:  Property Law
CC         STREET:  301 Henrietta Street
CC         CITY:  Kalamazoo
CC         STATE:  Michigan
CC         COUNTRY:  USA
CC         ZIP:  49001
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US94/06430
CC         FILING DATE:
CC         CLASSIFICATION:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Darnley Jr., James D.
CC         REGISTRATION NUMBER:  33,673
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  616-385-5210
CC         TELEFAX:  616-385-6897
CC         TELEX:  224401
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  249 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
SQ   SEQUENCE   249 AA; 27770 MW; 295210 CN;

  Query Match           49.5%;  Score 46;  DB 13;  Length 249;
  Best Local Similarity  50.0%;  Pred. No. 6.75e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db     82 INFMRDKDPN 91
          ||| |:::|:
Qy      1 INFTRQRQPS 10


RESULT   24
ID   US-08-479-638-3     STANDARD;     PRT;   303 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 3, Application US/08479638.
XX
CC    Sequence 3, Application US/08479638
CC    Patent No. 5695968
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nanba, Hirokazu
CC       APPLICANT:  Yamada, Yukio
CC       APPLICANT:  Takano, Masayuki
CC       APPLICANT:  Ikenaka, Yasuhiro
```

```
CC       APPLICANT:  Takahashi, Satomi
CC       APPLICANT:  Yajima, Kazuyoshi
CC       TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF D-ALPHA-AMINO
CC       TITLE OF INVENTION:  ACIDS
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Wegner, Cantor, Mueller & Player
CC         STREET:  1233 20th Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         ZIP:  20036-8218
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/479,638
CC         FILING DATE:  07-JUN-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/917,111
CC         FILING DATE:  07-AUG-1992
CC         APPLICATION NUMBER:  JP 400848/1990
CC         FILING DATE:  07-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 407922/1990
CC         FILING DATE:  27-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 078840/1991
CC         FILING DATE:  11-APR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/JP91/01696
CC         FILING DATE:  06-DEC-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Player Esq., William E.
CC         REGISTRATION NUMBER:  31,409
CC         REFERENCE/DOCKET NUMBER:  P-500-23486
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-887-0400
CC         TELEFAX:  202-835-0605
CC         TELEX:  440706
CC     INFORMATION FOR SEQ ID NO:  3:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  303 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Agrobacterium radiobacter
CC         STRAIN:  KNK 712 (FERM BP-1900)
SQ   SEQUENCE   303 AA; 34154 MW; 451356 CN;

  Query Match           49.5%;  Score 46;  DB 7;  Length 303;
  Best Local Similarity  55.6%;  Pred. No. 6.75e+01;
  Matches    5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    286 FNFKQHRQP 294
          :|| ::|||
Qy      1 INFTRQRQP 9


RESULT   25
ID   US-07-917-111-2     STANDARD;     PRT;   303 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/07917111.
XX
CC    Sequence 2, Application US/07917111
```

```
CC    Patent No. 5565344
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nanba, Hirokazu
CC       APPLICANT:  Yamada, Yukio
CC       APPLICANT:  Takano, Masayuki
CC       APPLICANT:  Ikenaka, Yasuhiro
CC       APPLICANT:  Takahashi, Satomi
CC       APPLICANT:  Yajima, Kazuyoshi
CC       TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF D-ALPHA-AMINO
CC       TITLE OF INVENTION:  ACIDS
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Wegner, Cantor, Mueller & Player
CC         STREET:  1233 20th Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         ZIP:  20036-8218
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/917,111
CC         FILING DATE:  19920807
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 400848/1990
CC         FILING DATE:  07-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 407922/1990
CC         FILING DATE:  27-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 078840/1991
CC         FILING DATE:  11-APR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/JP91/01696
CC         FILING DATE:  06-DEC-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Player Esq., William E.
CC         REGISTRATION NUMBER:  31,409
CC         REFERENCE/DOCKET NUMBER:  P-500-23486
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-887-0400
CC         TELEFAX:  202-835-0605
CC         TELEX:  440706
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  303 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   303 AA; 34154 MW; 451356 CN;

  Query Match           49.5%;  Score 46;  DB 6;  Length 303;
  Best Local Similarity  55.6%;  Pred. No. 6.75e+01;
  Matches    5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    286 FNFKQHRQP 294
          :|| ::|||
Qy      1 INFTRQRQP 9


RESULT   26
ID   US-07-917-111-3     STANDARD;     PRT;   303 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 3, Application US/07917111.
XX
CC    Sequence 3, Application US/07917111
```

```
CC    Patent No. 5565344
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nanba, Hirokazu
CC       APPLICANT:  Yamada, Yukio
CC       APPLICANT:  Takano, Masayuki
CC       APPLICANT:  Ikenaka, Yasuhiro
CC       APPLICANT:  Takahashi, Satomi
CC       APPLICANT:  Yajima, Kazuyoshi
CC       TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF D-ALPHA-AMINO
CC       TITLE OF INVENTION:  ACIDS
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Wegner, Cantor, Mueller & Player
CC         STREET:  1233 20th Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         ZIP:  20036-8218
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/917,111
CC         FILING DATE:  19920807
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 400848/1990
CC         FILING DATE:  07-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 407922/1990
CC         FILING DATE:  27-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 078840/1991
CC         FILING DATE:  11-APR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/JP91/01696
CC         FILING DATE:  06-DEC-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Player Esq., William E.
CC         REGISTRATION NUMBER:  31,409
CC         REFERENCE/DOCKET NUMBER:  P-500-23486
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-887-0400
CC         TELEFAX:  202-835-0605
CC         TELEX:  440706
CC     INFORMATION FOR SEQ ID NO:  3:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  303 amino acids
CC         TYPE:  AMINO ACID
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Agrobacterium radiobacter
CC         STRAIN:  KNK 712 (FERM BP-1900)
SQ   SEQUENCE   303 AA; 34154 MW; 451356 CN;

  Query Match           49.5%;  Score 46;  DB 6;  Length 303;
  Best Local Similarity  55.6%;  Pred. No. 6.75e+01;
  Matches    5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    286 FNFKQHRQP 294
          :|| ::|||
Qy      1 INFTRQRQP 9


RESULT   27
ID   US-08-479-638-2     STANDARD;     PRT;   303 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/08479638.
XX
CC    Sequence 2, Application US/08479638
CC    Patent No. 5695968
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nanba, Hirokazu
CC       APPLICANT:  Yamada, Yukio
CC       APPLICANT:  Takano, Masayuki
CC       APPLICANT:  Ikenaka, Yasuhiro
CC       APPLICANT:  Takahashi, Satomi
CC       APPLICANT:  Yajima, Kazuyoshi
CC       TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF D-ALPHA-AMINO
CC       TITLE OF INVENTION:  ACIDS
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Wegner, Cantor, Mueller & Player
CC         STREET:  1233 20th Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         ZIP:  20036-8218
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/479,638
CC         FILING DATE:  07-JUN-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/917,111
CC         FILING DATE:  07-AUG-1992
CC         APPLICATION NUMBER:  JP 400848/1990
CC         FILING DATE:  07-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 407922/1990
CC         FILING DATE:  27-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 078840/1991
CC         FILING DATE:  11-APR-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/JP91/01696
CC         FILING DATE:  06-DEC-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Player Esq., William E.
CC         REGISTRATION NUMBER:  31,409
CC         REFERENCE/DOCKET NUMBER:  P-500-23486
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-887-0400
CC         TELEFAX:  202-835-0605
CC         TELEX:  440706
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  303 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   303 AA; 34154 MW; 451356 CN;

  Query Match           49.5%;  Score 46;  DB 7;  Length 303;
  Best Local Similarity  55.6%;  Pred. No. 6.75e+01;
  Matches    5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    286 FNFKQHRQP 294
          :|| ::|||
Qy      1 INFTRQRQP 9


RESULT   28
ID   US-07-977-434-6     STANDARD;     PRT;   830 AA.
XX
AC   xxxxxx
```

```
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application US/07977434.
XX
CC    Sequence 6, Application US/07977434
CC    Patent No. 5466591
CC     GENERAL INFORMATION:
CC       APPLICANT:  Gelfand, David H.
CC       APPLICANT:  Abramson, Richard D.
CC       TITLE OF INVENTION:  5' TO 3' EXONUCLEASE MUTATIONS OF
CC       TITLE OF INVENTION:  THERMOSTABLE DNA POLYMERASES
CC       NUMBER OF SEQUENCES:  38
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Hoffmann-La Roche Inc.
CC         STREET:  340 Kingsland Street
CC         CITY:  Nutley
CC         STATE:  New Jersey
CC         ZIP:  07110-1199
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  Macintosh
CC         OPERATING SYSTEM:  7
CC         SOFTWARE:  WordPerfect 2.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/977,434
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,490
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,466
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,213
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 523,394
CC         FILING DATE:  15-MAY-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 143,441
CC         FILING DATE:  12-JAN-1988
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 063,509
CC         FILING DATE:  17-JUN-1987
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 899,241
CC         FILING DATE:  22-AUG-1986
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 746,121
CC         FILING DATE:  15-AUG-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  WO PCT/US90/07641
CC         FILING DATE:  21-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 585,471
CC         FILING DATE:  20-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 455,611
CC         FILING DATE:  22-DEC-1989
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 609,157
CC         FILING DATE:  02-NOV-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 557,517
CC         FILING DATE:  24-JUL-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Luann Cserr
CC         REGISTRATION NUMBER:  31,822
CC         REFERENCE/DOCKET NUMBER:  Case No. 5466591 8753
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (510) 814-2972
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  830 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   830 AA; 93347 MW; 3328578 CN;

  Query Match           49.5%;  Score 46;  DB 5;  Length 830;
  Best Local Similarity  33.3%;  Pred. No. 6.75e+01;
  Matches     3;  Conservative    6;  Mismatches   0;  Indels   0;  Gaps   0;

Db    254 VDFAKRREP 262
          ::|:::|:|
Qy      1 INFTRQRQP 9


RESULT   29
ID   PCT-US91-07035-6     STANDARD;     PRT;   830 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application PC/TUS9107035.
XX
CC    Sequence 6, Application PC/TUS9107035
CC     GENERAL INFORMATION:
CC       APPLICANT:  Gelfand, David H.
CC       APPLICANT:  Abramson, Richard D.
CC       TITLE OF INVENTION:  5' TO 3' EXONUCLEASE MUTATIONS OF
CC       TITLE OF INVENTION:  THERMOSTABLE DNA POLYMERASES
CC       NUMBER OF SEQUENCES:  38
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cetus Corporation
CC         STREET:  1400 Fifty-third Street
CC         CITY:  Emeryville
CC         STATE:  California
CC         ZIP:  94608
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  WordPerfect 5.0
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US91/07035
CC         FILING DATE:  19910930
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,490
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,466
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,213
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 523,394
CC         FILING DATE:  15-MAY-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 143,441
CC         FILING DATE:  12-JAN-1988
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 063,509
CC         FILING DATE:  17-JUN-1987
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 899,241
CC         FILING DATE:  22-AUG-1986
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 746,121
CC         FILING DATE:  15-AUG-1991
CC       PRIOR APPLICATION DATA:
```

```
CC         APPLICATION NUMBER:  WO PCT/US90/07641
CC         FILING DATE:  21-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 585,471
CC         FILING DATE:  20-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 455,611
CC         FILING DATE:  22-DEC-1989
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 609,157
CC         FILING DATE:  02-NOV-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 557,517
CC         FILING DATE:  24-JUL-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Sias Ph.D, Stacey R.
CC         REGISTRATION NUMBER:  32,630
CC         REFERENCE/DOCKET NUMBER:  Case No. 2580
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-420-3300
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  830 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   830 AA; 93347 MW; 3328578 CN;

  Query Match           49.5%;  Score 46;  DB 10;  Length 830;
  Best Local Similarity  33.3%;  Pred. No. 6.75e+01;
  Matches    3;  Conservative    6;  Mismatches   0;  Indels   0;  Gaps   0;

Db    254 VDFAKRREP 262
          ::|:::|:|
Qy      1 INFTRQRQP 9


RESULT   30
ID   US-07-977-434-2     STANDARD;     PRT;   832 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/07977434.
XX
CC    Sequence 2, Application US/07977434
CC    Patent No. 5466591
CC     GENERAL INFORMATION:
CC       APPLICANT:  Gelfand, David H.
CC       APPLICANT:  Abramson, Richard D.
CC       TITLE OF INVENTION:  5' TO 3' EXONUCLEASE MUTATIONS OF
CC       TITLE OF INVENTION:  THERMOSTABLE DNA POLYMERASES
CC       NUMBER OF SEQUENCES:  38
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Hoffmann-La Roche Inc.
CC         STREET:  340 Kingsland Street
CC         CITY:  Nutley
CC         STATE:  New Jersey
CC         ZIP:  07110-1199
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  Macintosh
CC         OPERATING SYSTEM:  7
CC         SOFTWARE:  WordPerfect 2.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/977,434
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,490
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,466
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 590,213
CC         FILING DATE:  28-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 523,394
CC         FILING DATE:  15-MAY-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 143,441
CC         FILING DATE:  12-JAN-1988
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 063,509
CC         FILING DATE:  17-JUN-1987
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 899,241
CC         FILING DATE:  22-AUG-1986
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 746,121
CC         FILING DATE:  15-AUG-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  WO PCT/US90/07641
CC         FILING DATE:  21-DEC-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 585,471
CC         FILING DATE:  20-SEP-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 455,611
CC         FILING DATE:  22-DEC-1989
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 609,157
CC         FILING DATE:  02-NOV-1990
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 557,517
CC         FILING DATE:  24-JUL-1990
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Luann Cserr
CC         REGISTRATION NUMBER:  31,822
CC         REFERENCE/DOCKET NUMBER:  Case No. 5466591 8753
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (510) 814-2972
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  832 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   832 AA; 93909 MW; 3296737 CN;

  Query Match           49.5%;  Score 46;  DB 5;  Length 832;
  Best Local Similarity  33.3%;  Pred. No. 6.75e+01;
  Matches    3;  Conservative    6;  Mismatches   0;  Indels   0;  Gaps   0;

Db    256 VDFAKRREP 264
          ::|:::|:|
Qy      1 INFTRQRQP 9


RESULT   31
ID   US-08-156-020-6     STANDARD;     PRT;   832 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application US/08156020.
XX
CC    Sequence 6, Application US/08156020
CC    Patent No. 5474920
CC     GENERAL INFORMATION:
CC       APPLICANT:  Moses M.D., Robb E.
CC       TITLE OF INVENTION:  Modified Thermo-Resistant DNA
CC       TITLE OF INVENTION:  Polymerases
```

```
CC       NUMBER OF SEQUENCES:  15
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Allegretti & Witcoff
CC         STREET:  10 South Wacker Drive
CC         CITY:  Chicago
CC         STATE:  IL
CC         COUNTRY:  USA
CC         ZIP:  60606
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  Apple Macintosh
CC         OPERATING SYSTEM:  Macintosh
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/156,020
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Greenfield Ph.D., Michael S.
CC         REGISTRATION NUMBER:  37,142
CC         REFERENCE/DOCKET NUMBER:  93,413
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (312)715-1000
CC         TELEFAX:  (312)715-1234
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  832 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   832 AA; 93806 MW; 3296848 CN;

  Query Match           49.5%;  Score 46;  DB 5;  Length 832;
  Best Local Similarity  33.3%;  Pred. No. 6.75e+01;
  Matches     3;  Conservative    6;  Mismatches   0;  Indels   0;  Gaps   0;

Db    256 VDFAKRREP 264
          ::|:::|:|
Qy      1 INFTRQRQP 9


RESULT   32
ID   US-08-156-020-8     STANDARD;     PRT;   832 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/08156020.
XX
CC    Sequence 8, Application US/08156020
CC    Patent No. 5474920
CC     GENERAL INFORMATION:
CC       APPLICANT:  Moses M.D., Robb E.
CC       TITLE OF INVENTION:  Modified Thermo-Resistant DNA
CC       TITLE OF INVENTION:  Polymerases
CC       NUMBER OF SEQUENCES:  15
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Allegretti & Witcoff
CC         STREET:  10 South Wacker Drive
CC         CITY:  Chicago
CC         STATE:  IL
CC         COUNTRY:  USA
CC         ZIP:  60606
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  Apple Macintosh
CC         OPERATING SYSTEM:  Macintosh
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/156,020
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Greenfield Ph.D., Michael S.
CC         REGISTRATION NUMBER:  37,142
CC         REFERENCE/DOCKET NUMBER:  93,413
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (312)715-1000
CC         TELEFAX:  (312)715-1234
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  832 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   832 AA; 93840 MW; 3297291 CN;

  Query Match           49.5%;  Score 46;  DB 5;  Length 832;
  Best Local Similarity  33.3%;  Pred. No. 6.75e+01;
  Matches     3;  Conservative    6;  Mismatches   0;  Indels   0;  Gaps   0;

Db    256 VDFAKRREP 264
          ::|:::|:|
Qy      1 INFTRQRQP 9


RESULT   33
ID   US-08-156-020-4     STANDARD;     PRT;   832 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/08156020.
XX
CC    Sequence 4, Application US/08156020
CC    Patent No. 5474920
CC     GENERAL INFORMATION:
CC       APPLICANT:  Moses M.D., Robb E.
CC       TITLE OF INVENTION:  Modified Thermo-Resistant DNA
CC       TITLE OF INVENTION:  Polymerases
CC       NUMBER OF SEQUENCES:  15
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Allegretti & Witcoff
CC         STREET:  10 South Wacker Drive
CC         CITY:  Chicago
CC         STATE:  IL
CC         COUNTRY:  USA
CC         ZIP:  60606
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  Apple Macintosh
CC         OPERATING SYSTEM:  Macintosh
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/156,020
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Greenfield Ph.D., Michael S.
CC         REGISTRATION NUMBER:  37,142
CC         REFERENCE/DOCKET NUMBER:  93,413
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (312)715-1000
CC         TELEFAX:  (312)715-1234
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  832 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   832 AA; 93859 MW; 3296773 CN;

  Query Match           49.5%;  Score 46;  DB 5;  Length 832;
  Best Local Similarity  33.3%;  Pred. No. 6.75e+01;
```

```
  Matches     3;  Conservative    6;  Mismatches   0;  Indels   0;  Gaps   0;

Db     256 VDFAKRREP 264
           ::|:::|:|
Qy       1 INFTRQRQP 9


RESULT   34
ID   US-08-073-384C-4     STANDARD;     PRT;   832 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/08073384C.
XX
CC    Sequence 4, Application US/08073384C
CC    Patent No. 5541311
CC     GENERAL INFORMATION:
CC       APPLICANT:  Dahlberg, James E.
CC       APPLICANT:  Lyamichev, Victor I.
CC       APPLICANT:  Brow, Mary Ann D.
CC       TITLE OF INVENTION:  SYNTHESIS-DEFICIENT THERMOSTABLE DNA
CC       TITLE OF INVENTION:  POLYMERASE
CC       NUMBER OF SEQUENCES:  29
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  HAVERSTOCK, MEDLEN & CARROLL
CC         STREET:  220 Montgomery Street, Suite 2200
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  United States of America
CC         ZIP:  94104
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/073,384C
CC         FILING DATE:  04-JUN-1993
CC         CLASSIFICATION:  536
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/986,330
CC         FILING DATE:  07-DEC-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Carroll, Peter G.
CC         REGISTRATION NUMBER:  32,837
CC         REFERENCE/DOCKET NUMBER:  FORS-00613
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415/705-8410
CC         TELEFAX:  415/397-8338
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  832 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   832 AA; 93909 MW; 3296737 CN;

  Query Match           49.5%;  Score 46;  DB 6;  Length 832;
  Best Local Similarity  33.3%;  Pred. No. 6.75e+01;
  Matches     3;  Conservative    6;  Mismatches   0;  Indels   0;  Gaps   0;

Db     256 VDFAKRREP 264
           ::|:::|:|
Qy       1 INFTRQRQP 9


RESULT   35
ID   US-08-254-359A-4     STANDARD;     PRT;   832 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/08254359A.
XX
CC    Sequence 4, Application US/08254359A
CC    Patent No. 5614402
CC     GENERAL INFORMATION:
CC       APPLICANT:  DAHLBERG, JAMES E.
CC       APPLICANT:  LYAMICHEV, VICTOR I.
CC       APPLICANT:  BROW, MARY ANN D.
CC       TITLE OF INVENTION:  5' NUCLEASES DERIVED FROM THERMOSTABLE
CC       TITLE OF INVENTION:  DNA POLYMERASE
CC       NUMBER OF SEQUENCES:  40
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  HAVERSTOCK, MEDLEN & CARROLL
CC         STREET:  220 MONTGOMERY STREET, SUITE 2200
CC         CITY:  SAN FRANCISCO
CC         STATE:  CALIFORNIA
CC         COUNTRY:  UNITED STATES OF AMERICA
CC         ZIP:  94104
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/254,359A
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/073,384
CC         FILING DATE:  06-JUN-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/986,330
CC         FILING DATE:  07-DEC-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  CARROLL, PETER G.
CC         REGISTRATION NUMBER:  32,837
CC         REFERENCE/DOCKET NUMBER:  FORS-01000
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 705-8410]
CC         TELEFAX:  (415) 397-8338
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  832 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   832 AA; 93909 MW; 3296737 CN;

  Query Match           49.5%;  Score 46;  DB 7;  Length 832;
  Best Local Similarity  33.3%;  Pred. No. 6.75e+01;
  Matches     3;  Conservative    6;  Mismatches   0;  Indels   0;  Gaps   0;

Db     256 VDFAKRREP 264
           ::|:::|:|
Qy       1 INFTRQRQP 9


Search completed: Tue Apr  7 08:43:40 1998
Job time : 10 secs.
```

****

(TM)

****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:39:04 1998; MasPar time 2.05 Seconds
115.783 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-2
Description: (1-14) from 5541104.pep
Perfect Score: 93
Sequence: 1 INFTRQRQPSEGSS 14

Scoring table: PAM 150
Gap 15

Searched: 111725 seqs, 16919825 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: a-geneseq30
1:a-geneseq1

Statistics: Mean 20.114; Variance 30.066; scale 0.669

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	93	100.0	40	1	R80618	Immunogenic peptide of	1.92e-09
2	93	100.0	335	1	R70909	Human melanoma antigen	1.92e-09
3	57	61.3	935	1	R50092	Humanised anti-CEA sFv	1.46e-01
4	55	59.1	122	1	P81863	Sequence encoded by LA	3.62e-01
5	53	57.0	35	1	R65121	MAGE 1 immunogenic pep	8.86e-01
6	53	57.0	122	1	R12258	HIV-1 strain OYI open	8.86e-01
7	53	57.0	368	1	R95054	TGF-a-DETA-DGAL4 multi	8.86e-01
8	53	57.0	440	1	R36807	Pseudomonas exotoxin d	8.86e-01
9	53	57.0	440	1	R32455	PE amino acids 2-414.	8.86e-01
10	53	57.0	445	1	R07054	PE40AB protein compris	8.86e-01
11	53	57.0	446	1	R20200	TGF-alpha-PE40Ab.	8.86e-01
12	53	57.0	446	1	R06993	PE40Ab protein compris	8.86e-01
13	53	57.0	446	1	R06449	TGF-alpha-PE40-Ab modi	8.86e-01
14	53	57.0	446	1	W19869	TGF-57-PE40 encoded by	8.86e-01
15	53	57.0	446	1	R06447	TGF-57-Pseudomonas exo	8.86e-01
16	53	57.0	446	1	W19871	TGF-alpha-PE40Ab.	8.86e-01
17	53	57.0	447	1	R95055	IL-2-DETA-DGAL4 multid	8.86e-01
18	53	57.0	452	1	R32454	PE(2-414)-Ma(57-68) hy	8.86e-01
19	53	57.0	452	1	R36806	PE domains I and II fu	8.86e-01
20	53	57.0	460	1	R99581	Heregulin/Pseudomonas	8.86e-01
21	53	57.0	488	1	R91735	Heregulin-PE40 HAR-TX	8.86e-01
22	53	57.0	489	1	W05137	TGF alpha-ETA fusion p	8.86e-01
23	53	57.0	522	1	R04934	Immunotoxin hybrid of	8.86e-01
24	53	57.0	556	1	R95053	scFv(FRP5)-DETA-DGAL4	8.86e-01
25	53	57.0	575	1	R04920	Immunoprotein PEX46.	8.86e-01
26	53	57.0	583	1	R04923	Immunoprotein TANG11.	8.86e-01
27	53	57.0	600	1	R04919	Immunoprotein PEX45.	8.86e-01
28	53	57.0	603	1	R04924	Immunoprotein TANG12.	8.86e-01
29	53	57.0	639	1	R40105	Pseudomonas exotoxin (	8.86e-01
30	53	57.0	639	1	R40106	Pseudomonas exotoxin (	8.86e-01
31	53	57.0	639	1	R40107	Pseudomonas exotoxin (	8.86e-01
32	53	57.0	639	1	R40113	Pseudomonas exotoxin (	8.86e-01
33	53	57.0	639	1	R40110	Pseudomonas exotoxin (	8.86e-01
34	53	57.0	639	1	R40108	Pseudomonas exotoxin (	8.86e-01
35	53	57.0	639	1	R40109	Pseudomonas exotoxin (	8.86e-01
36	53	57.0	639	1	R40112	Pseudomonas exotoxin (	8.86e-01
37	53	57.0	639	1	R40111	Pseudomonas exotoxin (	8.86e-01
38	53	57.0	639	1	R40104	Pseudomonas exotoxin (	8.86e-01
39	53	57.0	639	1	R40102	Pseudomonas exotoxin f	8.86e-01
40	53	57.0	640	1	R87738	Native pseudomonas exo	8.86e-01
41	53	57.0	663	1	R26982	(FRP5)-ETA fusion prot	8.86e-01
42	53	57.0	663	1	R26983	(FRP51)-ETA fusion pro	8.86e-01
43	53	57.0	677	1	W05135	scFv(225)-ETA fusion p	8.86e-01
44	53	57.0	677	1	W05136	scFv(FRP5)-ETA fusion	8.86e-01
45	53	57.0	686	1	R32456	PE with inactivated to	8.86e-01
46	53	57.0	686	1	R36808	Pseudomonas Exotoxin w	8.86e-01
47	53	57.0	691	1	R32453	PE(2-414)-M1(2-252) hy	8.86e-01
48	53	57.0	691	1	R36805	Pseudomonas exotoxin-i	8.86e-01
49	53	57.0	696	1	R36820	PE-Influenza A virus M	8.86e-01
50	53	57.0	696	1	R32468	BSPEM1c5aa fragment.	8.86e-01
51	53	57.0	701	1	R39573	Sequence of 741 sFV-PE	8.86e-01
52	53	57.0	702	1	R32457	PE having M1 residues	8.86e-01
53	53	57.0	702	1	R36809	Full-length PE with In	8.86e-01
54	53	57.0	725	1	W05138	scFv(FRP5)/TGF alpha-E	8.86e-01
55	53	57.0	780	1	R36810	Full-length PE with In	8.86e-01
56	53	57.0	780	1	R32458	PE having M1 residues	8.86e-01
57	53	57.0	918	1	W05139	scFv2(FRP5/225)-ETA (v	8.86e-01
58	53	57.0	918	1	W05143	scFv2(FRP5/FRP5)-ETA (	8.86e-01
59	53	57.0	918	1	W05140	scFv2(225/FRP5)-ETA.	8.86e-01
60	53	57.0	921	1	W05142	scFv2(FRP5/FRP5)-ETA (	8.86e-01
61	53	57.0	925	1	W05144	scFv2(FRP5/FRP5)-ETA (	8.86e-01
62	53	57.0	943	1	R32469	PE binding and translo	8.86e-01
63	53	57.0	943	1	R36821	PE binding/translocati	8.86e-01
64	53	57.0	963	1	R36822	PE binding/translocati	8.86e-01
65	53	57.0	963	1	R32470	PE binding and translo	8.86e-01
66	53	57.0	1046	1	W05141	scFv2(FRP5/225)-ETA (v	8.86e-01
67	52	55.9	410	1	R47236	Wild-type Feline Herpe	1.38e+00
68	51	54.8	658	1	R23162	Mutant thermostable DN	2.13e+00
69	51	54.8	706	1	R23161	Mutant thermostable DN	2.13e+00
70	51	54.8	784	1	R23160	Mutant thermostable DN	2.13e+00
71	51	54.8	815	1	R23159	Mutant thermostable DN	2.13e+00
72	51	54.8	860	1	R23158	Mutant thermostable DN	2.13e+00
73	50	53.8	122	1	R10174	Rap (R) protein encode	3.28e+00
74	50	53.8	122	1	P81856	Sequence encoded by LA	3.28e+00
75	50	53.8	337	1	R25359	KNK-003A.	3.28e+00
76	50	53.8	915	1	R65159	Potassium ion channel	3.28e+00
77	49	52.7	46	1	R81761	Vpr/Vpx motif-derived	5.04e+00
78	49	52.7	47	1	R81758	Vpr/Vpx motif-derived	5.04e+00
79	49	52.7	50	1	R48953	HIV-1 HXB2 transactiva	5.04e+00
80	49	52.7	50	1	R48957	HIV-1 pX NLVPR.	5.04e+00
81	49	52.7	52	1	R81759	Vpr/Vpx motif-derived	5.04e+00
82	49	52.7	122	1	R48963	HIV VPR.	5.04e+00
83	49	52.7	228	1	R25117	Non-A, Non-B Hepatitis	5.04e+00
84	49	52.7	433	1	P94259	Modifiied nitrilase ge	5.04e+00
85	49	52.7	613	1	P94239	Modified nitilase gene	5.04e+00
86	49	52.7	741	1	W11859	Rat-derived oxide squa	5.04e+00
87	48	51.6	381	1	R38297	dTDP-D-glucose synthas	7.70e+00
88	48	51.6	3593	1	R44431	eryA region polypeptid	7.70e+00
89	47	50.5	395	1	R14848	Protein associated wit	1.17e+01
90	47	50.5	395	1	W00089	RAS-related protein en	1.17e+01
91	47	50.5	656	1	R94536	TEC tyrosine kinase.	1.17e+01
92	46	49.5	329	1	R46247	Heat-stable carbamylas	1.77e+01
93	46	49.5	329	1	R46260	Improved Heat-stable c	1.77e+01
94	46	49.5	329	1	R46265	Improved Heat-stable c	1.77e+01
95	46	49.5	329	1	R46240	Heat-stable carbamylas	1.77e+01
96	46	49.5	329	1	R46245	Heat-stable carbamylas	1.77e+01

97	46	49.5	329	1	R46250	Heat-stable carbamylas	1.77e+01
98	46	49.5	329	1	R46269	Improved Heat-stable c	1.77e+01
99	46	49.5	329	1	R46249	Heat-stable carbamylas	1.77e+01
100	46	49.5	329	1	R46266	Improved Heat-stable c	1.77e+01

ALIGNMENTS

```
RESULT    1
ID   R80618 standard; Protein; 14 AA.
AC   R80618;
DT   28-FEB-1996  (first entry)
DE   Immunogenic peptide of tumour rejection antigen (MAGE-1).
KW   Tumour rejection antigen; MAGE-1; monoclonal antibody; MAb;
KW   diagnosis; immunoassay; cancer; immunogen; antisera.
OS   Homo sapiens.
PN   WO9520974-A1.
PD   10-AUG-1995.
PF   05-JAN-1995; U00095.
PR   01-FEB-1994; US-190411.
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
PA   (SLOK ) MEMORIAL SLOAN-KETTERING CANCER CENT.
PI   Boon-falleur T,  Chen Y,  Garin-chesa P,  Old LJ,  Rettig WJ;
PI   Stockert E,  Van der bruggen P;
DR   WPI; 95-283606/37.
PT   New monoclonal antibody binding specifically to MAGE-1 - useful for
PT   diagnosis and monitoring of cancer, also new hybridomas, recombinant
PT   MAGE-1 and immunogenic peptide(s)
PS   Claim 12; Page 20; 33pp; English.
CC   A monoclonal antibody directed against the tumour rejection antigen
CC   (MAGE-1) can be used to detect MAGE-1 in samples by standard
CC   immunoassay methods for diagnosis and monitoring of cancer etc. The
CC   monoclonal antibody is designated MA454 and is produced by the
CC   hybridoma deposited as ATCC HBB11540.  The monoclonal antibody is
CC   specific for MAGE-1, having no reactivity for MAGE-2 or MAGE-3.
CC   Peptide fragments of MAGE-1 (See R80618-20) may be useful as
CC   immunogens for production of the monoclonal antibody and antisera.
SQ   Sequence   14 AA;

  Query Match           100.0%;  Score 93;  DB 1;  Length 40;
  Best Local Similarity 100.0%;  Pred. No. 1.92e-09;
  Matches    14;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 INFTRQRQPSEGSS 40
          ||||||||||||||
Qy      1 INFTRQRQPSEGSS 14


RESULT    2
ID   R70909 standard; Protein; 309 AA.
AC   R70909;
DT   09-OCT-1995  (first entry)
DE   Human melanoma antigen MAGE-1.
KW   Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
KW   HLA-restricted cytotoxic T-lymphocyte activity.
OS   Homo sapiens.
PN   WO9504542-A.
PD   16-FEB-1995.
PF   02-AUG-1994; U08721.
PR   06-AUG-1993; US-103623.
PA   (CYTE-) CYTEL CORP.
PI   Fikes JD,  Livingston BD,  Sette AD,  Sidney JC;
DR   WPI; 95-090681/12.
DR   N-PSDB; Q85435.
PT   Human melanoma antigen, MAGE-1, peptide(s) - useful for
PT   stimulating immune response against melanoma
PS   Example 1; Fig 1; 59pp; English.
CC   Q85435 encodes R70909 human melanoma antigen MAGE-1, it was used
CC   to produce the C-terminal MAGE-1 peptides described in R70915 to
CC   R70969. These peptides are useful for defining epitopes that
CC   engender a HLA-restricted cytotoxic lymphocyte activity against
CC   MAGE-1 antigens. Compsns. containing these peptides can be
CC   administered, as a vaccine to patients susceptable to MAGE
CC   associated tumours, e.g. melanomas.
SQ   Sequence   309 AA;

  Query Match           100.0%;  Score 93;  DB 1;  Length 335;
  Best Local Similarity 100.0%;  Pred. No. 1.92e-09;
  Matches    14;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     94 INFTRQRQPSEGSS 107
          ||||||||||||||
Qy      1 INFTRQRQPSEGSS 14


RESULT    3
ID   R50092 standard; Protein; 909 AA.
AC   R50092;
DT   26-OCT-1994  (first entry)
DE   Humanised anti-CEA sFv fragment-human beta-glucuronidase fusion
DE   protein.
KW   Carcinoembryonic antigen; single chain variable region; sFv fragment;
KW   fusion gene; cancer treatment; targetted drug delivery; tumour.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1. .19
FT   /label= signal_peptide
FT   Protein         20. .909
FT   /label= fusion_protein
FT   /note= "humanised anti-CEA sFv fragment fused to
FT   human beta-glucuronidase"
PN   EP-590530-A.
PD   06-APR-1994.
PF   24-SEP-1993; 115418.
PR   02-OCT-1992; DE-233152.
PA   (BEHW ) BEHRINGWERKE AG.
PI   Bosslet K,  Czech J,  Gehrmann M,  Seemann G;
DR   WPI; 94-111012/14.
DR   N-PSDB; Q58896.
PT   New fusion protein contg. enzyme for prodrug activation - coupled
PT   to antigen binding component, esp. sFv antibody fragment, partic.
PT   for treatment of tumours
PS   Claim 13; Page 12-15; 35pp; German.
CC   The sequence R50092 comprises a humanised sFv-fragment against CEA
CC   fused to a human beta-glucuronidase. The fusion protein is
CC   useful for targetting beta-glucuronidase to cancer cells expressing
CC   CEA, where the enzyme is able to convert a prodrug into its active
CC   form. Any fusion protein not bound to tumour can be removed by
CC   internalisation via the mannose-6-phosphate and galactose receptors.
SQ   Sequence   909 AA;

  Query Match            61.3%;  Score 57;  DB 1;  Length 935;
  Best Local Similarity  63.6%;  Pred. No. 1.46e-01;
  Matches     7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    893 FTRQRQPKSAA 903
          |||||||  ::
Qy      3 FTRQRQPSEGS 13


RESULT    4
ID   P81863 standard; protein; 96 AA.
AC   P81863;
DT   16-DEC-1990  (first entry)
DE   Sequence encoded by LAV MA L R gene
KW   HIV; HTLV III; AIDS; diagnosis; vaccine; probe; hybridisation.
OS   Lymphadenopathy associated virus MA L.
PN   WO8707906-A.
PD   30-DEC-1987.
PF   22-JUN-1987; E00326.
PR   23-JUN-1986; EP-401380.
PA   (INSP) Inst Pasteur.
PI   Alizon M, Sonigo P, Wain-Hobson S, Montagnier L;
DR   WPI; 88-014396/02.
DR   N-PSDB; N80437.
```

```
PT   New variants of lymphadenopathy associated virus (LAV) -
PT   used for prodn. of DNA, antigens and antibodies used in
PT   diagnosis of AIDS and pre-AIDS
PS   Claim 8; Fig 8A-8I; 72pp; English.
CC   LAV EL I (N80436) and LAV MA L (N80437) were isolated from the peripheral
CC   blood lymphocytes of patients. Different AIDS virus isolates concerned
CC   are designated by 3 letters of the patients name. Stable probes including
CC   the DNA sequences can be used for detection of the new LAV viruses or
CC   related viruses or DNA proviruses in eg. biological samples. The proteins
CC   or peptides can be used for detection of antibodies induced in vivo and
CC   present in biological fluids. The DNA can also be used for the expression
CC   of LAV viral antigens for the prodn. of a vaccine against LAV. The
CC   polypeptides can also be used for the prodn. of antibodies for the
CC   detection of proteins related to the LAV viruses, partic. for diagnosis
CC   of AIDS or pre-AIDS.
SQ   Sequence   96 AA;

  Query Match           59.1%;  Score 55;  DB 1;  Length 122;
  Best Local Similarity  57.1%;  Pred. No. 3.62e-01;
  Matches     8;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    107 IGITRQRRARNGSS 120
          | :|||:: :|||
Qy      1 INFTRQRQPSEGSS 14


RESULT    5
ID   R65121 standard; peptide; 9 AA.
AC   R65121;
DT   09-OCT-1995  (first entry)
DE   MAGE 1 immunogenic peptide 66-74.
KW   MAGE 1; immunogenic peptide 66-74; cytotoxic C cells;
KW   in vitro activation; cancer; AIDS; bacterial infections; malaria;
KW   fungal infections; tuberculosis; hepatitis.
OS   Homo sapiens.
PN   WO9504817-A.
PD   16-FEB-1995.
PF   01-AUG-1994; U08672.
PR   06-AUG-1993; US-103401.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Kubo R,  Serra H,  Tsai V,  Wentworth P;
DR   WPI; 95-090895/12.
PT   In vitro activation of cytotoxic T cells for selected killing of
PT   target cells  - for treating e.g. cancer, AIDS, hepatitis etc.by
PT   incubating them with antigen presenting cells loaded with
PT   appropriate immunogenic peptide
PS   Example 3; Page 35; 53pp; English.
CC   R65109-R65145 are immunogenic peptides, they are used in a new
CC   method for the in vitro activation of cytotoxic T cells (CTC).
CC   This is achieved by incubating the CTCs with antigen presenting
CC   cells loaded with an appropriate immunogenic peptide (e.g. one
CC   of the above peptides). By selecting the peptides used the
CC   following diseases and infections can be treated; cancer, AIDS,
CC   hepatitis, other viral and bacterial infections, malaria and
CC   tuberculosis.
SQ   Sequence   9 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 8.86e-01;
  Matches     7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     29 INFTRQR 35
          |||||||
Qy      1 INFTRQR 7


RESULT    6
ID   R12258 standard; Protein; 96 AA.
AC   R12258;
DT   20-AUG-1991  (first entry)
DE   HIV-1 strain OYI open reading frame (ORF) R protein.
KW   HIV-1; AIDS; retroviruses.
OS   Homo sapiens.
PN   US5019510-A.
PD   28-MAY-1991.
PF   28-OCT-1987; 113655.
PR   28-OCT-1987; US-113655.
PA   (INSP ) INST PASTEUR.
PI   Wain-Hobson S, Huet T, Delaporte E, Brun-Vezinet F;
DR   WPI; 91-177518/24.
PT   Purified human retrovirus - is mutant of HIV-1 having
PT   characteristics of HIV-1 OYI, used in diagnosis of HIV infection
PS   Disclosure; fig 4; 23pp; English.
CC   This sequence constitutes the ORF R protein constituent of a new
CC   strain of HIV-1 retrovirus, OYI.  This mutant retroviral strain is
CC   useful in an assay for diagnosing HIV infection.  See also Q11943
CC   (OYI nucleotide sequence), R12255-57 and R12259-62 (other HIV OYI
CC   constituent proteins).
SQ   Sequence   96 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 122;
  Best Local Similarity  50.0%;  Pred. No. 8.86e-01;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db    107 IGITRQRRARNGAS 120
          | :|||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT    7
ID   R95054 standard; Protein; 342 AA.
AC   R95054;
DT   19-AUG-1996  (first entry)
DE   TGF-a-DETA-DGAL4 multidomain protein.
KW   Nucleic acid transfer system; gene transfer; gene therapy;
KW   cell targeting; multidomain protein; vector; cancer;
KW   exotoxin A; DETA; ompA; signal peptide; GAL4; TGF-a;
KW   transforming growth factor-alpha.
OS   Chimeric synthetic;
OS   Chimeric Homo sapiens;
OS   Chimeric Pseudomonas aeruginosa;
OS   Chimeric Saccharomyces cerevisiae.
FH   Key             Location/Qualifiers
FT   Peptide         1. .8
FT   /label= FLAG_epitope
FT   Peptide         9. .12
FT   /label= Spacer
FT   Domain          13. .62
FT   /label= TGF-a
FT   /note= "amino acids 1-50 of human TGF-a
FT   Peptide         63. .65
FT   /label= Spacer
FT   Peptide         66. .71
FT   /label= Hexa-histidine
FT   Peptide         72
FT   /label= Spacer
FT   Domain          73. .187
FT   /label= ETA
FT   /note= "amino acids 252-366 of ETA"
FT   Peptide         188
FT   /label= Spacer
FT   Domain          189. .334
FT   /label= GAL4
FT   /note= "amino acids 2-147 of yeast GAL4"
FT   Peptide         335. .342
FT   /label= Spacer
FT   /note= "endoplasmic reticulum retention signal"
PN   WO9613599-A1.
PD   09-MAY-1996.
PF   31-OCT-1995; E04270.
PR   01-NOV-1994; EP-810627.
PA   (WELS/) WELS W.
PI   Fominaya J,  Wels W;
DR   WPI; 96-239505/24.
DR   N-PSDB; T29409.
PT   Nucleic acid transfer system for gene therapy, e.g. against cancer
```

```
PT   - includes toxin translocation domain to target nucleic acid to
PT   specific cell
PS   Claim 7; Page 64-65; 106pp; English.
CC   A multidomain protein (R95054) has a FLAG epitope, a portion
CC   of human transforming growth factor-alpha (TGF-a) that acts as a
CC   ligand domain, a non-cytotoxic portion of Pseudomonas aeruginosa
CC   exotoxin A acting as a translocation domain and the DNA
CC   binding domain of yeast GAL4.  It is the product of a fusion
CC   gene (T29410) and can be expressed in E. coli (resulting in
CC   removal of an ompA signal peptide).  It is used with an effector
CC   nucleic acid that comprises e.g. a gene to be delivered to
CC   a cell and a cognate structure for the GAL4 DNA binding domain.
CC   This provides a novel means of nucleic acid transfer, suitable
CC   for gene therapy.
SQ   Sequence   342 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 368;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches    6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    119 FTRHRQP 125
          |||:|||
Qy      3 FTRQRQP 9


RESULT    8
ID   R36807 standard; Protein; 414 AA.
AC   R36807;
DT   25-AUG-1993  (first entry)
DE   Pseudomonas exotoxin domains I and II encoded by pVC-PEBT.
KW   Vaccine; cytotoxic T lymphocyte; CTL; influenza A virus;
KW   matrix protein; Ma; Pseudomonas exotoxin; cell recognition domain;
KW   translocation domain; anti-viral agent.
OS   Pseudomonas aeruginosa.
PN   EP-541335-A.
PD   12-MAY-1993.
PF   04-NOV-1992; 310067.
PR   08-NOV-1991; US-792507.
PA   (MERI ) MERCK & CO INC.
PI   Donnelly JJ,  Friedman A,  Hawe LA,  Liu MA,  Marshall MS;
PI   Montgomery DL,  Oliff AA,  Shi X,  Ulmer J;
DR   WPI; 93-154266/19.
DR   N-PSDB; Q41715.
PT   Recombinant DNA encoding bacterial toxin-antigen conjugates - are
PT   useful as vaccines against viral infections, tumours and
PT   parasites
PS   Example 5; Page 30-32; 81pp; English.
CC   Control plasmid pVC-PEBT encodes a T7 promoter-driven gene fusion
CC   consisting of PE amino acids 2-414 followed by termination codons,
CC   instead of by at least part of the influenza A virus Matrix
CC   protein (as in e.g. Q41714).
SQ   Sequence   414 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 440;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches    6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    299 FTRHRQP 305
          |||:|||
Qy      3 FTRQRQP 9


RESULT    9
ID   R32455 standard; Protein; 414 AA.
AC   R32455;
DT   20-JUL-1993  (first entry)
DE   PE amino acids 2-414.
KW   PE; Pseudomonas exotoxin; influenza A virus; M1; matrix protein;
KW   T7 polymerase; fusion; hybrid; pVC-PEBT; pVC-PEM1-2.
OS   Synthetic.
PN   EP-532090-A.
PD   17-MAR-1993.
PF   02-SEP-1992; 202660.
PR   09-SEP-1991; US-756249.
PA   (MERI ) MERCK & CO INC.
PI   Donnelly JJ,  Friedman A,  Hawe LA,  Liu MA,  Marshall MS;
PI   Montgomery DL,  Oliff AI,  Shi X,  Ulmer J;
DR   WPI; 93-087107/11.
DR   N-PSDB; Q37108.
PT   Bacterial toxin-antigen protein conjugates - to elicit cytotoxic
PT   T-lymphocyte immune response, used for preventing viral
PT   infections, e.g. by influenza virus, HIV and human
PT   papilloma:virus
PS   Disclosure; Page 33-35; 85pp; English.
CC   Example 5 describes the construction of pVC-PEBT.
CC   A control plasmid was constructed which encodes a T7 polymerase
CC   driven gene fusion consisting of PE amino acids 2 to 414 followed by
CC   termination codons.  pVC-PEM1-2 was digested with SacII and EcoRI to
CC   remove the M1 sequence.  The vector was gel purified and ligated to
CC   an oligonucleotide that builds back PE codon no. 414 followed by
CC   termination signals shown in Q37893.  The resulting construction
CC   was named pVC-PEBT (Q37108).
SQ   Sequence   414 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 440;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches    6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    299 FTRHRQP 305
          |||:|||
Qy      3 FTRQRQP 9


RESULT   10
ID   R07054 standard; protein; 419 AA.
AC   R07054;
DT   18-JAN-1991  (first entry)
DE   PE40AB protein comprising a portion of the Pseudomonas exotoxin A.
KW   TGF-alpha-PE40; PE40ab; tumour; epidermal growth factor; EGF;
KW   transforming growth factor-alpha; TGF-alpha.
OS   Pseudomonas sp.
PN   EP-389043-A.
PD   26-SEP-1990.
PF   15-MAR-1990; 200613.
PR   22-MAR-1989; US-327214.
PA   (MERI ) MERCK & CO INC.
PI   Riemen MW, Stirdivant SM;
DR   WPI; 90-291988/39.
DR   N-PSDB; Q06127.
PT   Modified PE40 by substitution with other amino acids for cysteine -
PT   improving specificity of targetting agent for tumour cells.
PS   Disclosure; Table 3; 21pp; English.
CC   By replacing cysteine residues at positions 265 and 287 and/or 372
CC   and 379, chemical ambiguities may be eliminated, and targeting
CC   specificity for targeted agents of tumour cells eg. EGF or TGF-alpha
CC   may be improved.
SQ   Sequence   419 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 445;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches    6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    105 FTRHRQP 111
          |||:|||
Qy      3 FTRQRQP 9


RESULT   11
ID   R20200 standard; Protein; 420 AA.
AC   R20200;
DT   16-APR-1992  (first entry)
DE   TGF-alpha-PE40Ab.
KW   Pseudomonas exotoxin; bladder; mutant; target; receptor binding.
FH   Key             Location/Qualifiers
FT   Region          5. .54
FT   /label= TGFalpha1-50
```

```
FT   Region          59. .420
FT   /label= PE252-613
FT   Misc_difference 176
FT   /note= "Ser -> Thr"
FT   Misc_difference 179
FT   /note= "Cys -> Ala"
FT   Misc_difference 186
FT   /note= "Cys -> Ala"
PN   EP-467536-A.
PD   22-JAN-1992.
PF   20-JUN-1991; 305582.
PR   21-JUN-1990; US-542281.
PR   14-MAR-1991; US-669269.
PA   (MERI ) MERCK & CO INC.
PI   Ahern J, Heimbrook DC, Oliff AI, Stirdivant SM;
DR   WPI; 92-026359/04.
PT   Treatment of bladder cancer using hybrid protein - comprising
PT   cell targetting agent e.g. EGF that binds to EGF receptor on
PT   tumour cells and PE40 cell toxin
PS   Disclosure; Page 18-19; 34pp; English.
CC   The modified PE40 domains of the hybrid proteins have two or four of
CC   the Cys residues (designated Cys265, Cys287, Cys372 and Cys372)
CC   substituted with neutral amino acids, e.g. Gly, Ala, or Phe.
CC   TGF-alpha-PE40aB (R20199) has Cys265 and Cys287 replaced;
CC   TGF-alpha-PE40Ab (R20200) has Cys372 and Cys379 replaced; and
CC   TGF-alpha-PE40ab (R20201) has all four replaced.
CC   The modified hybrid proteins were produced in E.coli transformed
CC   with TAC expression vectors.  Site specific mutations were introduced
CC   to the unmodified TGF-alpha-PE40 gene cloned in pTACTGF57-PE40.
CC   The mol. efficiently targets receptors on human bladder tumour cells
CC   (the modified PE40 domain has improved receptor binding) and is
CC   used for selectively killing bladder tumour cells.
SQ   Sequence   420 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 446;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    105 FTRHRQP 111
          |||:|||
Qy      3 FTRQRQP 9


RESULT   12
ID   R06993 standard; protein; 420 AA.
AC   R06993;
DT   18-JAN-1991  (first entry)
DE   PE40Ab protein comprising a portion of the Pseudomonas exotoxin A
DE   lacking cysteine residues at 372 and 379.
KW   TGF-alpha-PE40; PE40ab; tumour; epidermal growth factor; EGF;
KW   transforming growth factor-alpha; TGF-alpha.
OS   Pseudomonas sp.
PN   EP-389043-A.
PD   26-SEP-1990.
PF   15-MAR-1990; 200613.
PR   22-MAR-1989; US-327214.
PA   (MERI ) MERCK & CO INC.
PI   Riemen MW, Stirdivant SM;
DR   WPI; 90-291988/39.
PT   Modified PE40 by substitution with other amino acids for cysteine -
PT   improving specificity of targetting agent for tumour cells.
PS   Disclosure; Table 5; 21pp; English.
CC   By replacing cysteine residues at positions 265 and 287 and/or 372
CC   and 379, chemical ambiguities may be eliminated, and targeting
CC   specificity for targeted agents of tumour cells eg. EGF or TGF-alpha
CC   may be improved.
CC   See also Q06127.
SQ   Sequence   420 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 446;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    105 FTRHRQP 111
          |||:|||
Qy      3 FTRQRQP 9


RESULT   13
ID   R06449 standard; protein; 420 AA.
AC   R06449;
DT   04-JAN-1991  (first entry)
DE   TGF-alpha-PE40-Ab modified pseudomonas exotoxin hybrid protein.
KW   Pseudomonas exotoxin-40 (PE40); protein targeting agent;
KW   psoriasis treatment; anti-tumour agent; TGF-alpha-PE40-Ab;
FH   Key             Location/Qualifiers
FT   Region          1. .54
FT   /label=residues -4 to +50 of TGF-alpha
FT   Region          58. .420
FT   /label=residues +252 to +613 of PE
FT   /note="residues 369, 372 and 379 are modified"
PN   EP-383599-A.
PD   22-AUG-1990.
PF   15-FEB-1990; 301639.
PR   17-FEB-1989; US-312540.
PR   03-AUG-1989; US-389092.
PR   21-DEC-1989; US-449187.
PA   (MERI ) MERCK & CO INC.
PI   Oliff A, Jones DD, Edwards GM;.
DR   WPI; 90-255832/34.
PT   Modified pseudomonas exotoxin hybrid proteins - has at least 2
PT   cysteine residues replaced or deleted to improve binding to
PT   receptors.
PS   Example ; Table 5; 20pp; English.
CC   Modified pseudomonas exotoxin (PE40) linked to
CC   5' portion of transforming growth factor (TGF)-alpha as a targeting
CC   agent. Three site-specific mutations have been introduced c.f wild-
CC   type PE40. The Cys residues at positions 369 and 379 of the
CC   exotoxin have been replaced by Ala residues. Ser at position 369
CC   has been replaced with Thr. These changes improve receptor binding.
CC   The hybrid protein can bind and kill tumour cells or keratinocytes
CC   possessing TGF receptors for treatment of psoriasis or warts.
CC   See also R06447-R06448 and R06450
SQ   Sequence   420 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 446;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    105 FTRHRQP 111
          |||:|||
Qy      3 FTRQRQP 9


RESULT   14
ID   W19869 standard; Protein; 420 AA.
AC   W19869;
DT   19-AUG-1997  (first entry)
DE   TGF-57-PE40 encoded by pTAC TGF-57-PE40.
KW   Fusion protein; transforming growth factor-alpha; TGF-alpha; PE40;
KW   Pseudomonas exotoxin A; PE-A; translocation; enzymatic domain;
KW   cytosol; cellular intoxication; targeting domain; PE40aB; PE40ab;
KW   targeting-toxin molecule; target specificity; epidermal growth factor;
KW   EGF; cytotoxic; cancer.
OS   Chimeric - Homo sapiens.
OS   Chimeric - Pseudomonas aeruginosa.
FH   Key             Location/Qualifiers
FT   Protein         1. .58
FT   /label= TGF-alpha
FT   Protein         59. .420
FT   /label= PE40
FT   Misc_difference 143. .145
FT   /note= "encoded by GCCCTG"
FT   Misc_difference 72
FT   /note= "Corresponds to Cys265"
FT   Misc_difference 94
```

FT /note= "Corresponds to Cys287"
FT Misc_difference 179
FT /note= "Corresponds to Cys372"
FT Misc_difference 186
FT /note= "Corresponds to Cys379"
PN US5621078-A.
PD 15-APR-1997.
PF 22-MAR-1989; 327214.
PR 22-MAR-1989; US-327214.
PR 24-JUN-1991; US-708267.
PR 30-APR-1992; US-879037.
PR 10-SEP-1993; US-120698.
PR 21-FEB-1995; US-391259.
PA (MERI) MERCK & CO INC.
PI Riemen MW, Stirdivant SM;
DR WPI; 97-235227/21.
DR N-PSDB; T72116.
PT Modified Pseudomonas exotoxin 40 polypeptide(s) - for production of
PT cancer-specific cytotoxic conjugates
PS Example 1; Column 13-16; 15pp; English.
CC This sequence represents a fusion protein containing transforming growth
CC factor (TGF)-alpha and PE40. PE40 comprises residues 252-613 of
CC Pseudomonas exotoxin A (PE-A) and has a molecular weight of 40 kD.
CC PE40 retains the translocation and enzymatic domains responsible for
CC delivery of the protein to the cytosol and cellular intoxication
CC activity, but lacks the PE-A targeting domain. The invention concerns
CC modified PE40 polypeptides, PE40aB, which has Ala at positions 265 and
CC 287, and PE40ab, which has Ala at positions 265, 287, 372 and 379
CC (numbering according to the native 66kD PE-A sequence). The Cys
CC residues at positions 265, 287, 372 and 379 were found to interfere
CC with the construction of "targeting-toxin" molecules by chemical
CC conjugation methods. Substitution of the Cys residues in these positions
CC allows conjugates with greater target specificity to be prepared.
CC Conjugates comprising the modified PE40 peptides and epidermal growth
CC factor (EGF) or TGF-alpha (see also W19870-72) are cytotoxic and can
CC be targeted to cancers.
SQ Sequence 420 AA;

Query Match 57.0%; Score 53; DB 1; Length 446;
Best Local Similarity 85.7%; Pred. No. 8.86e-01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 105 FTRHRQP 111
||| :|||
Qy 3 FTRQRQP 9

RESULT 15
ID R06447 standard; protein; 420 AA.
AC R06447;
DT 04-JAN-1991 (first entry)
DE TGF-57-Pseudomonas exotoxin 40 fusion protein.
KW Pseudomonas exotoxin-40 (PE40); protein targeting agent;
KW psoriasis treatment; anti-tumour agent;
FH Key Location/Qualifiers
FT Region 1. .54
FT /label=residues -4 to +50 of TGF-alpha
FT Region 58. .420
FT /label=residues +252 to +613 of PE
PN EP-383599-A.
PD 22-AUG-1990.
PF 15-FEB-1990; 301639.
PR 17-FEB-1989; US-312540.
PR 03-AUG-1989; US-389092.
PR 21-DEC-1989; US-449187.
PA (MERI) MERCK & CO INC.
PI Oliff A, Jones DD, Edwards GM;.
DR WPI; 90-255832/34.
DR N-PSDB; Q05666.
PT Modified pseudomonas exotoxin hybrid proteins - has at least 2
PT cysteine residues replaced or deleted to improve binding to
PT receptors.
PS Example ; Table 2; 20pp; English.
CC Modified pseudomonas exotoxin (PE40) linked to
CC 5' portion of transforming growth factor (TGF)-alpha as a targeting
CC agent. The corresponding nucleotide sequence was constructed from
CC a synthetic oligonucleotide encoding the 5' portion of TGF-alpha
CC and linked to PE40 and a linker cassette called "cassette 57". The
CC recombinant plasmid was used to transform E.coli JM109 cells.
CC The hybrid protein can bind and kill tumour cells or keratinocytes
CC possessing TGF receptors for treatment of psoriasis or warts.
CC See also R06448-R06450
SQ Sequence 420 AA;

Query Match 57.0%; Score 53; DB 1; Length 446;
Best Local Similarity 85.7%; Pred. No. 8.86e-01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 105 FTRHRQP 111
||| :|||
Qy 3 FTRQRQP 9

RESULT 16
ID W19871 standard; Protein; 420 AA.
AC W19871;
DT 19-AUG-1997 (first entry)
DE TGF-alpha-PE40Ab.
KW Fusion protein; transforming growth factor-alpha; TGF-alpha; PE40;
KW Pseudomonas exotoxin A; PE-A; translocation; enzymatic domain;
KW cytosol; cellular intoxication; targeting domain; PE40aB; PE40ab;
KW targeting-toxin molecule; target specificity; epidermal growth factor;
KW EGF; cytotoxic; cancer.
OS Chimeric - Homo sapiens.
OS Chimeric - Pseudomonas aeruginosa.
OS Synthetic.
FH Key Location/Qualifiers
FT Protein 1. .58
FT /label= TGF-alpha
FT Protein 59. .420
FT /label= PE40Ab
FT Modified_site 72
FT /note= "Corresponds to Cys265"
FT Misc_difference 94
FT /note= "Corresponds to Cys287"
FT Misc_difference 179
FT /label= Cys372Ala
FT Misc_difference 186
FT /label= Cys379Ala
PN US5621078-A.
PD 15-APR-1997.
PF 22-MAR-1989; 327214.
PR 22-MAR-1989; US-327214.
PR 24-JUN-1991; US-708267.
PR 30-APR-1992; US-879037.
PR 10-SEP-1993; US-120698.
PR 21-FEB-1995; US-391259.
PA (MERI) MERCK & CO INC.
PI Riemen MW, Stirdivant SM;
DR WPI; 97-235227/21.
PT Modified Pseudomonas exotoxin 40 polypeptide(s) - for production of
PT cancer-specific cytotoxic conjugates
PS Example 2; Column 21-24; 15pp; English.
CC The sequences given in W19870-72 represent fusion proteins containing
CC transforming growth factor (TGF)-alpha and modified PE40 proteins.
CC PE40 comprises residues 252-613 of Pseudomonas exotoxin A (PE-A) and
CC has a molecular weight of 40 kD. PE40 retains the translocation and
CC enzymatic domains responsible for delivery of the protein to the cytosol
CC and cellular intoxication activity, but lacks the PE-A targeting domain.
CC The modified PE40 proteins are PE40aB, which has Ala at positions 265 and
CC 287, PE40Ab, which has Ala at positions 372 and 379, and PE40ab, which
CC has Ala at positions 265, 287, 372 and 379 (numbering according to the
CC native 66kD PE-A sequence). The Cys residues at positions 265, 287, 372
CC and 379 were found to interfere with the construction of "targeting-
CC toxin" molecules by chemical conjugation methods. Substitution of the
CC Cys residues in these positions allows conjugates with greater target

```
CC   specificity to be prepared.  The modified PE40 proteins are named
CC   according to which domain the modified Cys residues appear in.  Domain A
CC   comprises Cys 265 and 287, and domain B comprises Cys 372 and 379.
CC   An upper case letter (A, B) indicates that Cys is present in these
CC   positions, whereas a lower case letter (a, b) indicates that these
CC   positions have been mutated to Ala. Conjugates comprising the modified
CC   PE40 peptides and epidermal growth factor (EGF) or TGF-alpha are
CC   cytotoxic and can be targeted to cancers.
SQ   Sequence   420 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 446;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    105 FTRHRQP 111
          |||:|||
Qy      3 FTRQRQP 9


RESULT   17
ID   R95055 standard; Protein; 421 AA.
AC   R95055;
DT   19-AUG-1996  (first entry)
DE   IL-2-DETA-DGAL4 multidomain protein.
KW   Nucleic acid transfer system; gene transfer; gene therapy;
KW   cell targeting; multidomain protein; vector; cancer;
KW   exotoxin A; DETA; ompA; signal peptide; GAL4; interleukin-2;
KW   IL-2.
OS   Chimeric synthetic;
OS   Chimeric Homo sapiens;
OS   Chimeric Pseudomonas aeruginosa;
OS   Chimeric Saccharomyces cerevisiae.
FH   Key             Location/Qualifiers
FT   Peptide         1. .8
FT   /label= FLAG_epitope
FT   Peptide         9. .17
FT   /label= Spacer
FT   Domain          18. .150
FT   /label= IL-2
FT   /note= "amino acids 1-113 of human IL-2
FT   Peptide         151
FT   /label= Spacer
FT   Domain          152. .266
FT   /label= ETA
FT   /note= "amino acids 252-366 of ETA"
FT   Peptide         267
FT   /label= Spacer
FT   Domain          268. .413
FT   /label= GAL4
FT   /note= "amino acids 2-147 of yeast GAL4"
FT   Peptide         414. .421
FT   /label= Spacer
FT   /note= "endoplasmic reticulum retention signal"
PN   WO9613599-A1.
PD   09-MAY-1996.
PF   31-OCT-1995; E04270.
PR   01-NOV-1994; EP-810627.
PA   (WELS/) WELS W.
PI   Fominaya J,  Wels W;
DR   WPI; 96-239505/24.
DR   N-PSDB; T29411.
PT   Nucleic acid transfer system for gene therapy, e.g. against cancer
PT   - includes toxin translocation domain to target nucleic acid to
PT   specific cell
PS   Claim 7; Page 67-69; 106pp; English.
CC   A multidomain protein (R95055) has a FLAG epitope, a portion
CC   of human interleukin-2 that acts as a ligand domain, a
CC   non-cytotoxic portion of Pseudomonas aeruginosa exotoxin A acting
CC   as a translocation domain and the DNA binding domain of yeast GAL4.
CC   It is the product of a fusion gene (T29411) and can be expressed
CC   in E. coli (resulting in removal of an ompA signal peptide).  It is
CC   used with an effector nucleic acid that comprises e.g. a gene to be
CC   delivered to a cell and a cognate structure for the GAL4 DNA binding
CC   domain.  This provides a novel means of nucleic acid transfer,
CC   suitable for gene therapy.
SQ   Sequence   421 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 447;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    198 FTRHRQP 204
          |||:|||
Qy      3 FTRQRQP 9


RESULT   18
ID   R32454 standard; Protein; 426 AA.
AC   R32454;
DT   20-JUL-1993  (first entry)
DE   PE(2-414)-Ma(57-68) hybrid protein.
KW   PE; Pseudomonas exotoxin; influenza A virus; M1; matrix protein;
KW   ompA; leader; signal; fusion; hybrid; BS-PEMa-1; pVC45DR+T;
KW   pVC-ompA-PEMa-1; pVC-PEMa-1; T7 promoter; RBS; ribosome binding site;
KW   initiation sequence; build-back.
OS   Synthetic.
PN   EP-532090-A.
PD   17-MAR-1993.
PF   02-SEP-1992; 202660.
PR   09-SEP-1991; US-756249.
PA   (MERI ) MERCK & CO INC.
PI   Donnelly JJ,  Friedman A,  Hawe LA,  Liu MA,  Marshall MS;
PI   Montgomery DL,  Oliff AI,  Shi X,  Ulmer J;
DR   WPI; 93-087107/11.
DR   N-PSDB; Q37892.
PT   Bacterial toxin-antigen protein conjugates - to elicit cytotoxic
PT   T-lymphocyte immune response, used for preventing viral
PT   infections, e.g. by influenza virus, HIV and human
PT   papilloma:virus
PS   Disclosure; Page 30-31; 85pp; English.
CC   Example 4 describes the subcloning of PEMa from BS-PEMa-1 into
CC   pVC45DF+T.  The PEMa insert (Q37890) was prepd. by restricting
CC   BS-PEMa-1 with SacI and removing the 3' overhang by treatment with
CC   T4 DNA polymerase, then restricting with ApaI and gel purifying.
CC   pVC45DF+T was restricted with EcoRI and the 5' overhang filled in
CC   with Klenow enzyme treatment.  It was subsequently restricted with
CC   ApaI and gel purified.  The vector and fragment were ligated together,
CC   and the resulting construction was named pVC-ompA-PEMa-1.
CC   The ompA signal sequence was removed from pVC-ompA-PEMa-1 by
CC   digestion with XbaI and HindIII.  An oligonucleotide fragment contg.
CC   the T7 promoter, ribosome binding site, initiation sequence and a
CC   build-back of the 5' end of the PE coding region (Q37891) was
CC   ligated into the vector.  The resulting plasmid construct was named
CC   pVC-PEMa-1 and encodes a T7 polymerase-driven gene fusion consisting
CC   of PE amino acids 2 through 414 joined to influenza M1 amino acids
CC   57 to 68 (Ma) (Q37892).
SQ   Sequence   426 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 452;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    299 FTRHRQP 305
          |||:|||
Qy      3 FTRQRQP 9


RESULT   19
ID   R36806 standard; Protein; 426 AA.
AC   R36806;
DT   25-AUG-1993  (first entry)
DE   PE domains I and II fused to influenza A virus Ma.
KW   Vaccine; cytotoxic T lymphocyte; CTL; influenza A virus;
KW   matrix protein; Ma; Pseudomonas exotoxin; cell recognition domain;
KW   translocation domain; anti-viral agent; fusion construct.
OS   Chimeric Pseudomonas aeruginosa.
```

```
OS   Chimeric Influenza A virus.
FH   Key             Location/Qualifiers
FT   Region          2. .414
FT   /note= "PE domains I and II"
FT   Region          415. .426
FT   /note= "amino acids 57-68 of Influenza A Virus
FT   Matrix protein"
PN   EP-541335-A.
PD   12-MAY-1993.
PF   04-NOV-1992; 310067.
PR   08-NOV-1991; US-792507.
PA   (MERI ) MERCK & CO INC.
PI   Donnelly JJ,  Friedman A,  Hawe LA,  Liu MA,  Marshall MS;
PI   Montgomery DL,  Oliff AA,  Shi X,  Ulmer J;
DR   WPI; 93-154266/19.
DR   N-PSDB; Q41714.
PT   Recombinant DNA encoding bacterial toxin-antigen conjugates - are
PT   useful as vaccines against viral infections, tumours and
PT   parasites
PS   Example 4; Page 25; 81pp; English.
CC   An MI gene fragment (encoding amino acids 57-68 of influenza A virus
CC   matrix protein) was subcloned into BS-PE, a plasmid constructed by
CC   inserting a 1.3kb NruI/SacII fragment of plasmid pVC45-DF+T containing
CC   the domain I and II coding regions of Pseudomonas exotoxin into
CC   pBluescript II SK restricted with HincII and SacII. The PE-derived
CC   portion of the hybrid protein allows internalisation of the protein
CC   by an antigen-presenting cell. The hybrid protein is then processed
CC   and an antigenic segment (i.e. the Influenza A virus matrix protein)
CC   is presented on the cell surface where it elicits an immune response.
SQ   Sequence   426 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 452;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    299 FTRHRQP 305
          |||:|||
Qy      3 FTRQRQP 9


RESULT   20
ID   R99581 standard; protein; 434 AA.
AC   R99581;
DT   07-NOV-1996  (first entry)
DE   Heregulin/Pseudomonas ligand toxin.
KW   Heregulin; HRG beta2; Pseudomonas; exotoxin; ligand toxin; cancer;
KW   treatment; breast cancer; treatment; selective targetting; erbB3;
KW   erbB4.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Region          1. .8
FT   /label= FLAG peptide.
FT   Region          10. .70
FT   /label= HRG beta2 protein sequence.
FT   Region          71. .434
FT   /label= Pseudomonas PE40 exotoxin sequence.
PN   WO9616176-A1.
PD   30-MAY-1996.
PF   17-NOV-1995; AU0767.
PR   22-NOV-1994; AU-009598.
PA   (CRCB-) CRC BIOPHARMACEUTICAL RES PTY LTD.
PI   Daly RJ,  Fiddes RJ;
DR   WPI; 96-268612/27.
PT   Ligand-toxin, partic. useful for treating cancer cells - comprising
PT   a sequence homologous to the EGF-like domain of heregulin beta 2
PT   fused to a toxin
PS   Claim 4; Figure 1; 23pp; English.
CC   The ligand toxin consists of the epidermal growth factor-like domain
CC   of heregulin beta2 (HRG beta2) amino acids 177-237, fused to a
CC   modified Pseudomonas exotoxin (PE40).  The HRG beta2-PE40 ligand
CC   toxin was engineered and expressed using the pFlag prokaryotic
CC   expression system.  The ligand toxin can be used for treating a
CC   cancer in which there is an overexpression of erbB3 and/or erbB4.
CC   It selectively targets such cancer cells with no toxic effects on
CC   nearby benign cells.  The ligand toxin can be used in particular for
CC   the treatment of breast cancer.
SQ   Sequence   434 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 460;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    119 FTRHRQP 125
          |||:|||
Qy      3 FTRQRQP 9


RESULT   21
ID   R91735 standard; Protein; 462 AA.
AC   R91735;
DT   05-JUL-1996  (first entry)
DE   Heregulin-PE40 HAR-TX beta-2 chimeric protein.
KW   HER4/p180erbB4; HER4; receptor tyrosine kinase; HAR-TX beta-2;
KW   epidermal growth factor receptor; cancer; diagnosis; therapy;
KW   amphiregulin; heregulin; toxin; PE40; breast carcinoma.
OS   Chimeric Rattus sp.;
OS   Chimeric Pseudomonas sp.;
OS   Chimeric synthetic.
FH   Key             Location/Qualifiers
FT   Peptide          1. .35
FT   /label= Sig_peptide
FT   /note= "amphiregulin leader sequence"
FT   Protein          36. .95
FT   /label= Heregulin
FT   Peptide          96. .98
FT   /label= Linker
FT   Protein          99. .462
FT   /label= PE40
PN   WO9612019-A2.
PD   25-APR-1996.
PF   10-OCT-1995; U13524.
PR   14-OCT-1994; US-323442.
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
PI   Culouscou J,  Hellstrom I,  Hellstrom KE,  Plowman GD;
PI   Shoyab M,  Siegall C;
DR   WPI; 96-222005/22.
DR   N-PSDB; T18534.
PT   DNA encoding receptor tyrosine kinase, HER4 - related to human
PT   epidermal growth factor receptor, used for diagnosis and therapy of
PT   human cancers
PS   Claim 57; Page 156-158; 203pp; English.
CC   Cytotoxic fusion protein HAR-TX beta-2 (R91735) comprises a chimeric
CC   rat heregulin beta-2 ligand and Pseudomonas exotoxin PE40, joined to
CC   an amphiregulin leader sequence that facilitates purification.  The
CC   fusion protein is generated by expression of the corresponding
CC   chimeric coding sequence (see T18534).  The fusion protein
CC   specifically targets cells that express receptor tyrosine kinase
CC   HER4 (see R91733), e.g. prostate carcinoma, bladder carcinoma and
CC   breast carcinoma cells, allowing targeted cancer therapy.
SQ   Sequence   462 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 488;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    147 FTRHRQP 153
          |||:|||
Qy      3 FTRQRQP 9


RESULT   22
ID   W05137 standard; Protein; 463 AA.
AC   W05137;
DT   29-JAN-1997  (first entry)
DE   TGF alpha-ETA fusion protein.
KW   Single chain antibody; scFv; monoclonal antibody; MAb; TGF;
```

```
KW   transforming growth factor alpha; receptor; plasmid pSW202-TGF;
KW   cancer; therapy; antitumour; exotoxin A; ETA.
OS   Chimeric Homo sapiens;
OS   Chimeric Pseudomonas aeruginosa;
OS   Chimeric synthetic.
FH   Key             Location/Qualifiers
FT   Peptide         1. .21
FT   /label= Sig_peptide
FT   /note= "ompA signal peptide"
FT   Peptide         22. .38
FT   /label= Spacer
FT   Protein         39. .278
FT   /label= TGF-alpha
FT   Peptide         279. .289
FT   /label= Spacer
FT   Protein         290. .651
FT   /label= ETA
FT   /note= "exotoxin A amino acids 252-613"
PN   EP-739984-A1.
PD   30-OCT-1996.
PF   26-APR-1995; 106275.
PR   26-APR-1995; EP-106275.
PA   (SANT-) SAN TUMORFORSCHUNGS GMBH.
PI   Groner B,  Schmidt M,  Wels W;
DR   WPI; 96-478748/48.
DR   N-PSDB; T42037.
PT   Bivalent fusion proteins that bind epidermal growth factor receptor
PT   or analogues - and comprise at least two different cell surface
PT   binding domain(s), useful for tumour therapy
PS   Example 8; Page 23-24; 52pp; English.
CC   TGF alpha-ETA (W05137) comprises human transforming growth
CC   factor (TGF) alpha joined to the Pseudomonas aeruginosa cytotoxin,
CC   exotoxin A (ETA).  It is encoded by plasmid pSW202-TGF (see also
CC   T42037) obtd. by ligating TGF-alpha cDNA into plasmid pSW202-5 (see
CC   also T42036).  The construct can be used to produce novel bivalent
CC   fusion proteins (see also W05138-44) in bacterial host cells, for
CC   use as antitumour agents.
SQ   Sequence   463 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 489;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    148 FTRHRQP 154
          |||:|||
Qy      3 FTRQRQP 9


RESULT   23
ID   R04934 standard; protein; 496 AA.
AC   R04934;
DT   26-SEP-1990  (first entry)
DE   Immunotoxin hybrid of human interleukin-2 and Pseudomonas exotoxin.
KW   Pseudomonas exotoxin; interleukin-2; PE40; diabetes;
KW   arthritis; Crohn's disease; autoimmune disease.
OS   Synthetic.
PN   EP-369316-A.
PD   23-MAY-1990.
PF   9-NOV-1989; 120754.
PR   17-NOV-1988; US-272356; EP-120754.
PA   (HOFF) Hoffmann-La Roche.
PI   Grace JW;
DR   WPI; 90-157493/21.
PT   Immuno:toxin contg. natural human IL-2 sequence as
PT   N-terminal component -
PT   and Pseudomonas exotoxin region, and encoding DNA sequence
PT   vectors and transformed cells, for treating auto:immune disease.
PS   Disclosure; p; English.
CC   Hybrid enzyme is useful in treating autoimmune diseases, type I
CC   diabetes, rheumatoid arthritis, Crohn's disease, and in supressing
CC   allograft and organ transplant rejection.
SQ   Sequence   496 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 522;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    181 FTRHRQP 187
          |||:|||
Qy      3 FTRQRQP 9


RESULT   24
ID   R95053 standard; Protein; 530 AA.
AC   R95053;
DT   18-AUG-1996  (first entry)
DE   scFv(FRP5)-DETA-DGAL4 multidomain protein.
KW   Nucleic acid transfer system; gene transfer; gene therapy;
KW   cell targeting; multidomain protein; vector; cancer;
KW   exotoxin A; DETA; single chain antibody; scFv; Gal4.
OS   Chimeric synthetic;
OS   Chimeric Mus sp.;
OS   Chimeric Pseudomonas aeruginosa;
OS   Chimeric Saccharomyces cerevisiae.
FH   Key             Location/Qualifiers
FT   Peptide         1. .8
FT   /label= FLAG_epitope
FT   Peptide         9. .17
FT   /label= Spacer
FT   Domain          18. .257
FT   /label= ScFv(FRP5)
FT   Peptide         258. .260
FT   /label= Spacer
FT   Domain          261. .375
FT   /label= ETA
FT   /note= "amino acids 252-366 of exotoxin-A"
FT   Peptide         376
FT   /label= Spacer
FT   Domain          377. .522
FT   /label= GAL4
FT   /note= "amino acids 2-147 of yeast GAL4"
FT   Peptide         523. .530
FT   /label= Spacer
FT   /note= "endoplasmic reticulum retention peptide"
PN   WO9613599-A1.
PD   09-MAY-1996.
PF   31-OCT-1995; E04270.
PR   01-NOV-1994; EP-810627.
PA   (WELS/) WELS W.
PI   Fominaya J,  Wels W;
DR   WPI; 96-239505/24.
DR   N-PSDB; T29409.
PT   Nucleic acid transfer system for gene therapy, e.g. against cancer
PT   - includes toxin translocation domain to target nucleic acid to
PT   specific cell
PS   Claim 7; Page 59-61; 106pp; English.
CC   A multidomain protein (R95053) has a FLAG epitope, a single
CC   chain antibody, scFv, of monoclonal antibody FRP5 (raised
CC   against human tumour cell HER2 antigen) that acts as a ligand
CC   domain, a non-cytotoxic portion of Pseudomonas aeruginosa
CC   exotoxin A acting as a translocation domain and the DNA
CC   binding domain of yeast GAL4.  It is the product of a fusion
CC   gene (T29409) and can be expressed in E. coli (resulting in
CC   removal of ompA signal peptide).  It is used with an effector
CC   nucleic acid that comprises e.g. a gene to be delivered to
CC   a cell and a cognate structure for the GAL4 DNA binding domain.
CC   This provides a novel means of nucleic acid transfer, suitable
CC   for gene therapy.
SQ   Sequence   530 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 556;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    307 FTRHRQP 313
          |||:|||
```

```
Qy      3 FTRQRQP 9


RESULT   25
ID   R04920 standard; protein; 549 AA.
AC   R04920;
DT   02-OCT-1990  (first entry)
DE   Immunoprotein PEX46.
KW   Soluble T4 protein; immunotoxin; Pseudomona endotoxin A; AIDS; HIV; ARC;
KW   angiogenin; fusion protein; PEX46; pEX46.
FH   Key             Location/Qualifiers
FT   Region          2. .182
FT   /label=AAs 3-183 of T4 protein
FT   Region          183. .549
FT   /label=AAs 253-613 of Pseudomonas exotoxin A (PE40)
PN   WO9004414-A.
PD   03-MAY-1990.
PF   18-OCT-1989; U04584.
PR   18-OCT-1988; US-259355.
PA   (BIOJ) Biogen Inc.
PI   Meade HM, Lobb RR, Gates LL, Winkler G;
DR   WPI; 90-163876/21.
PT   New immunotoxin contg. soluble T4 protein components and toxin -
PT   esp. Pseudomonas endotoxin A, for treating or controlling AIDS
PT   and related conditions, and new DNA sequences.
PS   Disclosure; 7pp; English.
CC   This fusion immunoprotein was produced by constructing a hybrid DNA
CC   sequence of a soluble T4 protein and a fragment of Pseudomonas exotoxin A
CC   (PE40) in which all binding region of PE was removed and which did not
CC   contain the boundary region.  The hybrid DNA can then be inserted into an
CC   expression vector and used to produce recombinant fusion protein which is
CC   useful for preventing or treating AIDS, ARC, and HIV infections.   The T4
CC   protein is an HIV receptor which binds to the virus or to infected cells
CC   carrying the gp120/160 marker antigen, so provides v. specific targetting
CC   with minimal damage to non-target cells. Unmodified chain of PEx45 has
CC   a Mol. Wt. of 59,658.
SQ   Sequence   549 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 575;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    234 FTRHRQP 240
          |||:|||
Qy      3 FTRQRQP 9


RESULT   26
ID   R04923 standard; protein; 557 AA.
AC   R04923;
DT   02-OCT-1990  (first entry)
DE   Immunoprotein TANG11.
KW   Soluble T4 protein; immunotoxin; Pseudomona endotoxin A; AIDS; HIV; ARC;
KW   angiogenin; fusion protein; TANG11.
FH   Key             Location/Qualifiers
FT   Region          1. .182
FT   /label=AAs 3-183 of T4 protein
FT   Region          184. .300
FT   /label=translocation domain of PE
FT   Region          301. .340
FT   /label=binding domain of PE
FT   Region          341. .429
FT   /label=linker from ADP ribosylation domain of PE
FT   Region          433. .559
FT   /label=mature angiogenin
PN   WO9004414-A.
PD   03-MAY-1990.
PF   18-OCT-1989; U04584.
PR   18-OCT-1988; US-259355.
PA   (BIOJ) Biogen Inc.
PA   Meade HM, Lobb RR, Gates LL, Winkler G;
DR   WPI;, 90-163876/21.
PT   New immunotoxin contg. soluble T4 protein components and toxin -
PT   esp. Pseudomonas endotoxin A, for treating or controlling AIDS
PT   and related conditions, and new DNA sequences.
PS   Disclosure; 7pp; English.
CC   This fusion immunoprotein was produced by constructing a hybrid DNA
CC   sequence of a soluble T4 protein, a fragment of Pseudomonas exotoxin A
CC   including the translocation domain and part of the ADP ribosylation
CC   domain, and the gene for mature angiogenin.  The hybrid DNA can then be
CC   inserted into an expression vector and used to produce recombinant fusion
CC   protein which is useful for preventing or treating AIDS, ARC, and HIV
CC   infections.   The T4 protein is an HIV receptor which binds to the virus
CC   or to infected cells carrying the gp120/160 marker Ag, so provides v.
CC   specific targetting  with minimal damage to non-target cells.
SQ   Sequence   557 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 583;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    234 FTRHRQP 240
          |||:|||
Qy      3 FTRQRQP 9


RESULT   27
ID   R04919 standard; protein; 574 AA.
AC   R04919;
DT   02-OCT-1990  (first entry)
DE   Immunoprotein PEX45.
KW   Soluble T4 protein; immunotoxin; Pseudomona endotoxin A; AIDS; HIV; ARC;
KW   angiogenin; fusion protein; PEX45; pEX45.
FH   Key             Location/Qualifiers
FT   Region          2. .182
FT   /label=AAs 3-183 of T4 protein
FT   Region          183. .574
FT   /label=AAs224-613 of Pseudomonas exotoxin A
PN   WO9004414-A.
PD   03-MAY-1990.
PF   18-OCT-1989; U04584.
PR   18-OCT-1988; US-259355.
PA   (BIOJ) Biogen Inc.
PI   Meade HM, Lobb RR, Gates LL, Winkler G;
DR   WPI; 90-163876/21.
PT   New immunotoxin contg. soluble T4 protein components and toxin -
PT   esp. Pseudomonas endotoxin A, for treating or controlling AIDS
PT   and related conditions, and new DNA sequences.
PS   Disclosure; 7pp; English.
CC   This fusion immunoprotein was produced by constructing a hybrid DNA
CC   sequence of a soluble T4 protein and a fragment of Pseudomonas exotoxin A
CC   contg. the translocation and ADP-ribosylation domains.  The hybrid DNA
CC   can then be inserted into an expression vector and used to produce
CC   recombinant fusion protein, useful for preventing or treating AIDS,
CC   ARC, and HIV infections. The T4 protein is an HIV receptor which binds to
CC   the virus or to infected cells carrying the gp120/160 marker antigen, so
CC   provides v. specific targetting with minimal damage to non-target cells.
CC   The unmodified chain of PEX45 has a Mol. Wt. of 62,491.
SQ   Sequence   574 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 600;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    259 FTRHRQP 265
          |||:|||
Qy      3 FTRQRQP 9


RESULT   28
ID   R04924 standard; protein; 577 AA.
AC   R04924;
DT   02-OCT-1990  (first entry)
DE   Immunoprotein TANG12.
KW   Soluble T4 protein; immunotoxin; Pseudomona endotoxin A; AIDS; HIV; ARC;
KW   angiogenin; fusion protein; TANG12.
```

```
FH   Key             Location/Qualifiers
FT   Region          1. .182
FT   /label=AAs 3-183 of T4 protein
FT   Region          184. .300
FT   /label=translocation domain of PE
FT   Region          301. .340
FT   /label=binding domain of PE
FT   Region          341. .429
FT   /label=linker from ADP ribosylation domain of PE
FT   Peptide         431. .454
FT   /label=angiogenin signal sequence
FT   Protein         455. .577
FT   /label=mature angiogenin
PN   WO9004414-A.
PD   03-MAY-1990.
PF   18-OCT-1989; U04584.
PR   18-OCT-1988; US-259355.
PA   (BIOJ) Biogen Inc.
PI   Meade HM, Lobb RR, Gates LL, Winkler G;
DR   WPI; 90-163876/21.
PT   New immunotoxin contg. soluble T4 protein components and toxin -
PT   esp. Pseudomonas endotoxin A, for treating or controlling AIDS
PT   and related conditions, and new DNA sequences.
PS   Disclosure; 7pp; English.
CC   This fusion immunoprotein was produced by constructing a hybrid DNA
CC   sequence of a soluble T4 protein, a fragment of Pseudomonas exotoxin A
CC   including the translocation domain and part of the ADP ribosylation
CC   domain, and the gene for the signal sequence and mature angiogenin.
CC   The hybrid DNA can then be inserted into an expression vector and used
CC   to produce recombinant fusion protein which is useful for preventing or
CC   treating AIDS, ARC, and HIV infections. The T4 protein is an HIV receptor
CC   which binds to the virus or infected cells carrying the gp120/160 marker
CC   antigen, so provides v. specific targetting with minimal damage to
CC   non-target cells.
SQ   Sequence   577 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 603;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    234 FTRHRQP 240
          |||:|||
Qy      3 FTRQRQP 9


RESULT   29
ID   R40105 standard; Protein; 613 AA.
AC   R40105;
DT   27-JAN-1994  (first entry)
DE   Pseudomonas exotoxin (S25C).
KW   Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
KW   target site; cytotoxin; unpaired cysteine; receptor; binding site;
KW   monoclonal antibody; ligand; cell surface; mutation;
KW   steric unpaired cysteine; s.u.c.
OS   Pseudomonas aeruginosa.
FH   Key             Location/Qualifiers
FT   Misc_difference 25
FT   /note= "unpaired cysteine residue replaces Ser"
PN   WO9315113-A.
PD   05-AUG-1993.
PF   15-JAN-1993; U00358.
PR   24-JAN-1992; US-825396.
PA   (TANO-) TANOX BIOSYSTEMS INC.
PI   Chang TW;
DR   WPI; 93-258616/32.
PT   Site-specifically mutated cytotoxin(s) with an unpaired cysteine
PT   - such that conjugation of a binding mol. to the Cys blocks
PT   receptor binding used as immuno:toxins for highly specific
PT   targetting
PS   Claim 3; Page 20-23; 30pp; English.
CC   The new mutated toxin has an unpaired cysteine residue in
CC   or near the cytotoxin's receptor-binding site, and retains the
CC   same receptor-binding ability and cytotoxicity as the native
CC   cytotoxins provided they are not conjugated with a binding mol.
CC   The toxins are cross-linked through the free SH group of their
CC   unpaired cysteine residues to binding mols. (including monoclonal
CC   antibodies, fragments and other ligands) to form immunotoxins, and
CC   these immunotoxins do not bind to the cell surface receptors which
CC   are bound by the native cytotoxin.  However, when the cross-linker
CC   is cleaved and the binding mol. is released, the cytotoxin regains
CC   its receptor-binding ability and its cytotoxicity.
SQ   Sequence   613 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 639;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    298 FTRHRQP 304
          |||:|||
Qy      3 FTRQRQP 9


RESULT   30
ID   R40106 standard; Protein; 613 AA.
AC   R40106;
DT   27-JAN-1994  (first entry)
DE   Pseudomonas exotoxin (S88C).
KW   Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
KW   target site; cytotoxin; unpaired cysteine; receptor; binding site;
KW   monoclonal antibody; ligand; cell surface; mutation;
KW   steric unpaired cysteine; s.u.c.
OS   Pseudomonas aeruginosa.
FH   Key             Location/Qualifiers
FT   Misc_difference 88
FT   /note= "unpaired cysteine residue replaces Ser"
PN   WO9315113-A.
PD   05-AUG-1993.
PF   15-JAN-1993; U00358.
PR   24-JAN-1992; US-825396.
PA   (TANO-) TANOX BIOSYSTEMS INC.
PI   Chang TW;
DR   WPI; 93-258616/32.
PT   Site-specifically mutated cytotoxin(s) with an unpaired cysteine
PT   - such that conjugation of a binding mol. to the Cys blocks
PT   receptor binding used as immuno:toxins for highly specific
PT   targetting
PS   Claim 3; Page 20-23; 30pp; English.
CC   The new mutated toxin has an unpaired cysteine residue in
CC   or near the cytotoxin's receptor-binding site, and retains the
CC   same receptor-binding ability and cytotoxicity as the native
CC   cytotoxins provided they are not conjugated with a binding mol.
CC   The toxins are cross-linked through the free SH group of their
CC   unpaired cysteine residues to binding mols. (including monoclonal
CC   antibodies, fragments and other ligands) to form immunotoxins, and
CC   these immunotoxins do not bind to the cell surface receptors which
CC   are bound by the native cytotoxin.  However, when the cross-linker
CC   is cleaved and the binding mol. is released, the cytotoxin regains
CC   its receptor-binding ability and its cytotoxicity.
SQ   Sequence   613 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 639;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    298 FTRHRQP 304
          |||:|||
Qy      3 FTRQRQP 9


RESULT   31
ID   R40107 standard; Protein; 613 AA.
AC   R40107;
DT   27-JAN-1994  (first entry)
DE   Pseudomonas exotoxin (S96C).
KW   Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
KW   target site; cytotoxin; unpaired cysteine; receptor; binding site;
```

```
KW   monoclonal antibody; ligand; cell surface; mutation;
KW   steric unpaired cysteine; s.u.c.
OS   Pseudomonas aeruginosa.
FH   Key             Location/Qualifiers
FT   Misc_difference 96
FT   /note= "unpaired cysteine residue replaces Ser"
PN   WO9315113-A.
PD   05-AUG-1993.
PF   15-JAN-1993; U00358.
PR   24-JAN-1992; US-825396.
PA   (TANO-) TANOX BIOSYSTEMS INC.
PI   Chang TW;
DR   WPI; 93-258616/32.
PT   Site-specifically mutated cytotoxin(s) with an unpaired cysteine
PT   - such that conjugation of a binding mol. to the Cys blocks
PT   receptor binding used as immuno:toxins for highly specific
PT   targetting
PS   Claim 3; Page 20-23; 30pp; English.
CC   The new mutated toxin has an unpaired cysteine residue in
CC   or near the cytotoxin's receptor-binding site, and retains the
CC   same receptor-binding ability and cytotoxicity as the native
CC   cytotoxins provided they are not conjugated with a binding mol.
CC   The toxins are cross-linked through the free SH group of their
CC   unpaired cysteine residues to binding mols. (including monoclonal
CC   antibodies, fragments and other ligands) to form immunotoxins, and
CC   these immunotoxins do not bind to the cell surface receptors which
CC   are bound by the native cytotoxin.  However, when the cross-linker
CC   is cleaved and the binding mol. is released, the cytotoxin regains
CC   its receptor-binding ability and its cytotoxicity.
SQ   Sequence   613 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 639;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    298 FTRHRQP 304
          |||:|||
Qy      3 FTRQRQP 9


RESULT   32
ID   R40113 standard; Protein; 613 AA.
AC   R40113;
DT   27-JAN-1994  (first entry)
DE   Pseudomonas exotoxin (S245C).
KW   Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
KW   target site; cytotoxin; unpaired cysteine; receptor; binding site;
KW   monoclonal antibody; ligand; cell surface; mutation;
KW   steric unpaired cysteine; s.u.c.
OS   Pseudomonas aeruginosa.
FH   Key             Location/Qualifiers
FT   Misc_difference 245
FT   /note= "unpaired cysteine residue replaces Ser"
PN   WO9315113-A.
PD   05-AUG-1993.
PF   15-JAN-1993; U00358.
PR   24-JAN-1992; US-825396.
PA   (TANO-) TANOX BIOSYSTEMS INC.
PI   Chang TW;
DR   WPI; 93-258616/32.
PT   Site-specifically mutated cytotoxin(s) with an unpaired cysteine
PT   - such that conjugation of a binding mol. to the Cys blocks
PT   receptor binding used as immuno:toxins for highly specific
PT   targetting
PS   Claim 3; Page 20-23; 30pp; English.
CC   The new mutated toxin has an unpaired cysteine residue in
CC   or near the cytotoxin's receptor-binding site, and retains the
CC   same receptor-binding ability and cytotoxicity as the native
CC   cytotoxins provided they are not conjugated with a binding mol.
CC   The toxins are cross-linked through the free SH group of their
CC   unpaired cysteine residues to binding mols. (including monoclonal
CC   antibodies, fragments and other ligands) to form immunotoxins, and
CC   these immunotoxins do not bind to the cell surface receptors which
CC   are bound by the native cytotoxin.  However, when the cross-linker
CC   is cleaved and the binding mol. is released, the cytotoxin regains
CC   its receptor-binding ability and its cytotoxicity.
SQ   Sequence   613 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 639;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    298 FTRHRQP 304
          |||:|||
Qy      3 FTRQRQP 9


RESULT   33
ID   R40110 standard; Protein; 613 AA.
AC   R40110;
DT   27-JAN-1994  (first entry)
DE   Pseudomonas exotoxin (S188C).
KW   Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
KW   target site; cytotoxin; unpaired cysteine; receptor; binding site;
KW   monoclonal antibody; ligand; cell surface; mutation;
KW   steric unpaired cysteine; s.u.c.
OS   Pseudomonas aeruginosa.
FH   Key             Location/Qualifiers
FT   Misc_difference 188
FT   /note= "unpaired cysteine residue replaces Ser"
PN   WO9315113-A.
PD   05-AUG-1993.
PF   15-JAN-1993; U00358.
PR   24-JAN-1992; US-825396.
PA   (TANO-) TANOX BIOSYSTEMS INC.
PI   Chang TW;
DR   WPI; 93-258616/32.
PT   Site-specifically mutated cytotoxin(s) with an unpaired cysteine
PT   - such that conjugation of a binding mol. to the Cys blocks
PT   receptor binding used as immuno:toxins for highly specific
PT   targetting
PS   Claim 3; Page 20-23; 30pp; English.
CC   The new mutated toxin has an unpaired cysteine residue in
CC   or near the cytotoxin's receptor-binding site, and retains the
CC   same receptor-binding ability and cytotoxicity as the native
CC   cytotoxins provided they are not conjugated with a binding mol.
CC   The toxins are cross-linked through the free SH group of their
CC   unpaired cysteine residues to binding mols. (including monoclonal
CC   antibodies, fragments and other ligands) to form immunotoxins, and
CC   these immunotoxins do not bind to the cell surface receptors which
CC   are bound by the native cytotoxin.  However, when the cross-linker
CC   is cleaved and the binding mol. is released, the cytotoxin regains
CC   its receptor-binding ability and its cytotoxicity.
SQ   Sequence   613 AA;

  Query Match           57.0%;  Score 53;  DB 1;  Length 639;
  Best Local Similarity  85.7%;  Pred. No. 8.86e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    298 FTRHRQP 304
          |||:|||
Qy      3 FTRQRQP 9


RESULT   34
ID   R40108 standard; Protein; 613 AA.
AC   R40108;
DT   27-JAN-1994  (first entry)
DE   Pseudomonas exotoxin (S158C).
KW   Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
KW   target site; cytotoxin; unpaired cysteine; receptor; binding site;
KW   monoclonal antibody; ligand; cell surface; mutation;
KW   steric unpaired cysteine; s.u.c.
OS   Pseudomonas aeruginosa.
FH   Key             Location/Qualifiers
FT   Misc_difference 158
```

FT /note= "unpaired cysteine residue replaces Ser"
PN WO9315113-A.
PD 05-AUG-1993.
PF 15-JAN-1993; U00358.
PR 24-JAN-1992; US-825396.
PA (TANO-) TANOX BIOSYSTEMS INC.
PI Chang TW;
DR WPI; 93-258616/32.
PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
PT - such that conjugation of a binding mol. to the Cys blocks
PT receptor binding used as immuno:toxins for highly specific
PT targetting
PS Claim 3; Page 20-23; 30pp; English.
CC The new mutated toxin has an unpaired cysteine residue in
CC or near the cytotoxin's receptor-binding site, and retains the
CC same receptor-binding ability and cytotoxicity as the native
CC cytotoxins provided they are not conjugated with a binding mol.
CC The toxins are cross-linked through the free SH group of their
CC unpaired cysteine residues to binding mols. (including monoclonal
CC antibodies, fragments and other ligands) to form immunotoxins, and
CC these immunotoxins do not bind to the cell surface receptors which
CC are bound by the native cytotoxin. However, when the cross-linker
CC is cleaved and the binding mol. is released, the cytotoxin regains
CC its receptor-binding ability and its cytotoxicity.
SQ Sequence 613 AA;

Query Match 57.0%; Score 53; DB 1; Length 639;
Best Local Similarity 85.7%; Pred. No. 8.86e-01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 298 FTRHRQP 304
||| :|||
Qy 3 FTRQRQP 9

RESULT 35
ID R40109 standard; Protein; 613 AA.
AC R40109;
DT 27-JAN-1994 (first entry)
DE Pseudomonas exotoxin (R182C).
KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
KW monoclonal antibody; ligand; cell surface; mutation;
KW steric unpaired cysteine; s.u.c.
OS Pseudomonas aeruginosa.
FH Key Location/Qualifiers
FT Misc_difference 182
FT /note= "unpaired cysteine residue replaces Arg"
PN WO9315113-A.
PD 05-AUG-1993.
PF 15-JAN-1993; U00358.
PR 24-JAN-1992; US-825396.
PA (TANO-) TANOX BIOSYSTEMS INC.
PI Chang TW;
DR WPI; 93-258616/32.
PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
PT - such that conjugation of a binding mol. to the Cys blocks
PT receptor binding used as immuno:toxins for highly specific
PT targetting
PS Claim 3; Page 20-23; 30pp; English.
CC The new mutated toxin has an unpaired cysteine residue in
CC or near the cytotoxin's receptor-binding site, and retains the
CC same receptor-binding ability and cytotoxicity as the native
CC cytotoxins provided they are not conjugated with a binding mol.
CC The toxins are cross-linked through the free SH group of their
CC unpaired cysteine residues to binding mols. (including monoclonal
CC antibodies, fragments and other ligands) to form immunotoxins, and
CC these immunotoxins do not bind to the cell surface receptors which
CC are bound by the native cytotoxin. However, when the cross-linker
CC is cleaved and the binding mol. is released, the cytotoxin regains
CC its receptor-binding ability and its cytotoxicity.
SQ Sequence 613 AA;

Query Match 57.0%; Score 53; DB 1; Length 639;
Best Local Similarity 85.7%; Pred. No. 8.86e-01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 298 FTRHRQP 304
||| :|||
Qy 3 FTRQRQP 9

Search completed: Tue Apr 7 08:39:14 1998
Job time : 10 secs.

****

****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:38:32 1998; MasPar time 4.81 Seconds
124.671 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-2
Description: (1-14) from 5541104.pep
Perfect Score: 93
Sequence: 1 INFTRQRQPSEGSS 14

Scoring table: PAM 150
Gap 15

Searched: 195121 seqs, 42852602 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: pir55
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 22.202; Variance 34.214; scale 0.649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	93	100.0	280	2	JC2358	TOIG of: jc2358 check	6.37e-08
2	60	64.5	114	2	MADELEHQLQ	This is a DE line.	4.18e-01
3	58	62.4	172	2	I40349	TOIG of: i40349 check	9.87e-01
4	57	61.3	317	2	I38661	A;Accession: PH1297.	1.51e+00
5	57	61.3	317	2	JC2359	A;Accession: PH1298.	1.51e+00
6	57	61.3	651	2	A26581	A;Title: Mucopolysacch	1.51e+00
7	56	60.2	648	2	A29977	C;Superfamily: beta-gl	2.30e+00
8	56	60.2	648	2	I49692	TOIG of: i49692 check	2.30e+00
9	56	60.2	648	2	A32576	TOIG of: a32576 check	2.30e+00
10	56	60.2	648	2	B32576	TOIG of: b32576 check	2.30e+00
11	55	59.1	96	2	S54380	TOIG of: s54380 check	3.49e+00
12	55	59.1	1703	2	S15047	A;Accession: S16820.	3.49e+00
13	54	58.1	186	2	S46450	This is a DE line.	5.27e+00
14	54	58.1	355	2	S60396	A;Cross-references: EM	5.27e+00
15	54	58.1	373	2	A69158	A;Accession: A69158.	5.27e+00
16	53	57.0	167	2	B53293	A;Accession: B53293.	7.93e+00
17	53	57.0	188	1	THE	A;Title: The complete	7.93e+00
18	53	57.0	271	2	F69442	A;Accession: F69442.	7.93e+00
19	53	57.0	638	2	A30347	A;Accession: A30347.	7.93e+00
20	50	53.8	317	2	S24721	A;Accession: S24721.	2.63e+01
21	50	53.8	317	2	S24747	A;Accession: S24747.	2.63e+01
22	50	53.8	317	2	S24732	A;Accession: S24732.	2.63e+01
23	50	53.8	330	2	S10256	A;Accession: S10256.	2.63e+01
24	50	53.8	330	2	B60816	TOIG of: b60816 check	2.63e+01
25	50	53.8	410	2	A48585	TOIG of: a48585 check	2.63e+01
26	50	53.8	793	1	SURFCA	TOIG of: surfca check	2.63e+01
27	50	53.8	889	2	A41259	TOIG of: a41259 check	2.63e+01
28	49	52.7	102	1	DING	A;Accession: D26306.	3.89e+01
29	49	52.7	102	1	DING	A;Accession: B26306.	3.89e+01
30	49	52.7	128	1	NRYY	A;Accession: A00824.	3.89e+01
31	49	52.7	130	2	S03523	TOIG of: s03523 check	3.89e+01
32	49	52.7	237	2	A48912	TOIG of: a48912 check	3.89e+01
33	49	52.7	314	2	I68889	A;Accession: PH1294.	3.89e+01
34	49	52.7	343	2	S74937	A;Accession: S74937.	3.89e+01
35	49	52.7	349	2	A28658	A;Accession: A28658.	3.89e+01
36	49	52.7	364	1	AL	A;Accession: A93972.	3.89e+01
37	49	52.7	494	2	A47059	TOIG of: a47059 check	3.89e+01
38	49	52.7	542	2	JQ1524	TOIG of: jq1524 check	3.89e+01
39	49	52.7	648	2	A25047	A;Accession: A25047.	3.89e+01
40	49	52.7	933	2	S53523	F;316-933/Domain: form	3.89e+01
41	48	51.6	94	1	R3PJ19	TOIG of: r3pj19 check	5.71e+01
42	48	51.6	162	2	S75317	A;Accession: S75317.	5.71e+01
43	48	51.6	165	2	G64568	A;Authors: Hayes, W.S.	5.71e+01
44	48	51.6	179	1	E	A;Accession: A01107.	5.71e+01
45	48	51.6	220	2	S35789	TOIG of: s35789 check	5.71e+01
46	48	51.6	272	2	C32058	A;Accession: C32058.	5.71e+01
47	48	51.6	518	2	S42387	TOIG of: s42387 check	5.71e+01
48	48	51.6	551	3	S03667	TOIG of: s03667 check	5.71e+01
49	48	51.6	552	1	WJFFEN	A;Accession: A93354.	5.71e+01
50	48	51.6	559	2	I49444	TOIG of: i49444 check	5.71e+01
51	48	51.6	584	2	B25682	TOIG of: b25682 check	5.71e+01
52	48	51.6	3573	2	S23070	F;561-843/Domain: [acy	5.71e+01
53	47	50.5	159	2	A48428	TOIG of: a48428 check	8.36e+01
54	47	50.5	162	2	S76806	A;Accession: S76806.	8.36e+01
55	47	50.5	178	2	JC2022	A;Accession: JC2022.	8.36e+01
56	47	50.5	199	2	D69901	A;Status: preliminary;	8.36e+01
57	47	50.5	231	1	RDNCUF	This is a DE line.	8.36e+01
58	47	50.5	288	2	H64647	A;Authors: Hayes, W.S.	8.36e+01
59	47	50.5	292	1	A40650	A;Experimental source:	8.36e+01
60	47	50.5	297	2	G69892	A;Status: preliminary;	8.36e+01
61	47	50.5	305	2	S22313	TOIG of: s22313 check	8.36e+01
62	47	50.5	317	2	A40498	TOIG of: a40498 check	8.36e+01
63	47	50.5	321	2	S22822	TOIG of: s22822 check	8.36e+01
64	47	50.5	321	2	A48136	TOIG of: a48136 check	8.36e+01
65	47	50.5	322	2	A23677	A;Accession: A23677.	8.36e+01
66	47	50.5	322	2	I52910	TOIG of: i52910 check	8.36e+01
67	47	50.5	322	2	I58195	TOIG of: i58195 check	8.36e+01
68	47	50.5	324	2	JQ2292	This is a DE line.	8.36e+01
69	47	50.5	324	2	I39382	A;Accession: I39382.	8.36e+01
70	47	50.5	324	2	A55971	A;Accession: A55971.	8.36e+01
71	47	50.5	348	2	A49594	A;Cross-references: GB	8.36e+01
72	47	50.5	364	2	PS0015	TOIG of: ps0015 check	8.36e+01
73	47	50.5	397	2	S64841	TOIG of: s64841 check	8.36e+01
74	47	50.5	459	2	S03116	R;Lee, K.L.; Makkinje,	8.36e+01
75	47	50.5	460	2	B45353	TOIG of: b45353 check	8.36e+01
76	47	50.5	506	2	D69963	A;Experimental source:	8.36e+01
77	47	50.5	517	2	S20248	TOIG of: s20248 check	8.36e+01
78	47	50.5	527	2	S13763	TOIG of: s13763 check	8.36e+01
79	47	50.5	602	2	JU0215	TOIG of: ju0215 check	8.36e+01
80	47	50.5	608	2	JU0227	TOIG of: ju0227 check	8.36e+01
81	47	50.5	630	2	JU0228	TOIG of: ju0228 check	8.36e+01
82	47	50.5	631	2	I56997	TOIG of: i56997 check	8.36e+01
83	47	50.5	958	2	S47179	TOIG of: s47179 check	8.36e+01
84	47	50.5	984	1	GLYCEROL-3	A;Title: The complete	8.36e+01
85	47	50.5	1060	2	S54779	A;Cross-references: EM	8.36e+01
86	47	50.5	1359	2	S49883	A;Accession: S28406.	8.36e+01
87	46	49.5	152	1	NRRT	A;Accession: A91313.	1.22e+02
88	46	49.5	301	1	E	A;Title: The complete	1.22e+02
89	46	49.5	303	2	A38274	TOIG of: a38274 check	1.22e+02
90	46	49.5	320	1	D49349	A;Cross-references: EM	1.22e+02
91	46	49.5	328	2	S64306	TOIG of: s64306 check	1.22e+02
92	46	49.5	406	2	S17712	TOIG of: s17712 check	1.22e+02
93	46	49.5	425	1	KRXL2A	F;1-16/Region: coil 1A	1.22e+02
94	46	49.5	428	2	S07916	A;Accession: S07916.	1.22e+02
95	46	49.5	492	2	S71112	TOIG of: s71112 check	1.22e+02
96	46	49.5	492	2	S07124	TOIG of: s07124 check	1.22e+02

```
    97     46   49.5    504  2  S46752      TOIG of: s46752  check  1.22e+02
    98     46   49.5    831  3  JX0359      TOIG of: jx0359  check  1.22e+02
    99     46   49.5   1204  2  S62506      TOIG of: s62506  check  1.22e+02
   100     46   49.5   1507  2  B47328      TOIG of: b47328  check  1.22e+02


                              ALIGNMENTS

RESULT    1
ID   JC2358     STANDARD;     PRT;   280 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: jc2358  check: 467  from: 1  to: 280.
XX
CC     TOIG of: jc2358  check: 467  from: 1  to: 280
CC
CC    >P1;JC2358
CC    tumor-associated antigen , MAGE-1 - human
CC    C;Species: Homo sapiens (man)
CC    C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Mar-19
  96
CC    C;Accession: JC2358
CC    R;Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
CC    Biochem. Biophys. Res. Commun. 202, 549-555, 1994
CC    A;Title: Cloning and analysis of MAGE-1-related genes.
CC    A;Reference number: JC2358
CC    A;Accession: JC2358
CC    A;Molecule type: mRNA
CC    A;Residues: 1-280 <DIN>
CC    A;Experimental source: melanoma cell line DM150
CC    C;Genetics:
CC    A;Gene: MAGE
CC    F;161-169/Region: HLA-A1 binding #status predicted
SQ   SEQUENCE   280 AA; 30932 MW; 426797 CN;

  Query Match          100.0%;  Score 93;  DB 2;  Length 280;
  Best Local Similarity 100.0%;  Pred. No. 6.37e-08;
  Matches    14;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     68 INFTRQRQPSEGSS 81
          ||||||||||||||
Qy      1 INFTRQRQPSEGSS 14


RESULT    2
ID   MADELEHQLQTISLITVLGLDVKAELEAALLDYRDDLDIWKSHGYPVADLDQTVTVDKLLYMYMDAATAD
   STANDARD;     PRT;   114 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   This is a DE line.
XX
SQ   SEQUENCE   114 AA; 12387 MW; 67681 CN;

  Query Match           64.5%;  Score 60;  DB 2;  Length 114;
  Best Local Similarity  72.7%;  Pred. No. 4.18e-01;
  Matches     8;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db     87 FTRQRQPQDQS 97
          ||||||| : |
Qy      3 FTRQRQPSEGS 13


RESULT    3
ID   I40349     STANDARD;     PRT;   172 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: i40349  check: 9523  from: 1  to: 172.
XX
CC     TOIG of: i40349  check: 9523  from: 1  to: 172
CC
CC    >F1;I40349
CC    ribosomal protein L10 - Brucella abortus (fragment)
CC    C;Species: Brucella abortus
CC    C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Aug-19
  96
CC    C;Accession: I40349
CC    R;Oliveira, S.C.; Zhu, Y.; Splitter, G.A.
CC    Gene 140, 137-138, 1994
CC    A;Title: Sequences of the rplJL operon containing the L10 and L7/L12 gene
  s from Brucella abortus.
CC    A;Reference number: I40348; MUID:94171071
CC    A;Accession: I40349
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-172 <RES>
CC    A;Cross-references: GB:L23505; NID:g387910; PID:g387911
CC    C;Genetics:
CC    A;Gene: rplJ
SQ   SEQUENCE   172 AA; 18441 MW; 149510 CN;

  Query Match           62.4%;  Score 58;  DB 2;  Length 172;
  Best Local Similarity  60.0%;  Pred. No. 9.87e-01;
  Matches     6;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db     13 VKFVRQRQPG 22
          ::| |||||:
Qy      1 INFTRQRQPS 10


RESULT    4
ID   I38661     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: PH1297.
XX
CC    A;Accession: PH1297
CC    A;Molecule type: DNA
CC    A;Residues: 169-177 <TRA>
CC    C;Genetics:
CC    A;Gene: GDB:MAGE4
CC    A;Cross-references: GDB:331119
CC    A;Map position: Xq28-Xq28
CC    A;Introns: #status absent
SQ   SEQUENCE   317 AA; 34899 MW; 528124 CN;

  Query Match           61.3%;  Score 57;  DB 2;  Length 317;
  Best Local Similarity  71.4%;  Pred. No. 1.51e+00;
  Matches    10;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db     76 ISFTCWRQPNEGSS 89
          |:||  |||:||||
Qy      1 INFTRQRQPSEGSS 14


RESULT    5
ID   JC2359     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: PH1298.
XX
CC    A;Accession: PH1298
```

```
CC    A;Molecule type: DNA
CC    A;Residues: 169-177 <TRA>
CC    R;Fenton, R.G.
CC    submitted to the EMBL Data Library, June 1994
CC    A;Reference number: G07128
CC    A;Accession: G01446
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: mRNA
CC    A;Residues: 1-317 <FEN>
CC    A;Cross-references: EMBL:U10340; NID:g499123; PID:g499124
CC    C;Genetics:
CC    A;Gene: MAGE-X2
CC    F;169-177/Region: HLA-A1 binding #status predicted
SQ   SEQUENCE   317 AA; 34928 MW; 530585 CN;

  Query Match           61.3%;  Score 57;  DB 2;  Length 317;
  Best Local Similarity  71.4%;  Pred. No. 1.51e+00;
  Matches    10;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db     76 ISFTCWRQPNEGSS 89
          |:||  |||:||||
Qy      1 INFTRQRQPSEGSS 14


RESULT    6
ID   A26581     STANDARD;     PRT;   651 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: Mucopolysaccharidosis type VII: characterization of mutations and
   molecular heterogeneity.
XX
CC    A;Title: Mucopolysaccharidosis type VII: characterization of mutations an
  d molecular heterogeneity.
CC    A;Reference number: A36538; MUID:91090114
CC    A;Accession: A36538
CC    A;Molecule type: mRNA
CC    A;Residues: 378-385,616-621,643-651 <TOM>
CC    C;Genetics:
CC    A;Gene: GDB:GUSB
CC    A;Cross-references: GDB:120025; OMIM:253220
CC    A;Map position: 7q22-7q22
CC    C;Superfamily: beta-glucuronidase
CC    C;Keywords: glycoprotein; glycosidase; homotetramer; hydrolase; lysosome
CC    F;1-22/Domain: signal sequence #status predicted <SIG>
CC    F;23-651/Product: beta-glucuronidase, placental #status predicted <MAT>
SQ   SEQUENCE   651 AA; 74715 MW; 2313538 CN;

  Query Match           61.3%;  Score 57;  DB 2;  Length 651;
  Best Local Similarity  63.6%;  Pred. No. 1.51e+00;
  Matches     7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    609 FTRQRQPKSAA 619
          |||||||  ::
Qy      3 FTRQRQPSEGS 13


RESULT    7
ID   A29977     STANDARD;     PRT;   648 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   C;Superfamily: beta-glucuronidase.
XX
CC    C;Superfamily: beta-glucuronidase
CC    C;Keywords: glycosidase; hydrolase
CC    F;1-22/Domain: signal sequence #status predicted <SIG>
CC    F;23-648/Product: beta-glucuronidase #status predicted <MAT>
SQ   SEQUENCE   648 AA; 74239 MW; 2300715 CN;

  Query Match           60.2%;  Score 56;  DB 2;  Length 648;
  Best Local Similarity 100.0%;  Pred. No. 2.30e+00;
  Matches     7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    605 FTRQRQP 611
          |||||||
Qy      3 FTRQRQP 9


RESULT    8
ID   I49692     STANDARD;     PRT;   648 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: i49692  check: 9973  from: 1  to: 648.
XX
CC     TOIG of: i49692  check: 9973  from: 1  to: 648
CC
CC    >P1;I49692
CC    beta-glucuronidase precursor - mouse
CC    C;Species: Mus musculus (house mouse)
CC    C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Feb-19
  97
CC    C;Accession: I49692
CC    R;Funkenstein, B.; Leary, S.L.; Stein, J.C.; Catterall, J.F.
CC    Mol. Cell. Biol. 8, 1160-1168, 1988
CC    A;Title: Genomic organization and sequence of the Gus-s-a allele of the m
  urine beta-glucuronidase gene.
CC    A;Reference number: I49692; MUID:88216590
CC    A;Accession: I49692
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: mRNA
CC    A;Residues: 1-648 <RES>
CC    A;Cross-references: GB:M19279; NID:g193524; PID:g309257
CC    C;Superfamily: beta-glucuronidase
SQ   SEQUENCE   648 AA; 74192 MW; 2302638 CN;

  Query Match           60.2%;  Score 56;  DB 2;  Length 648;
  Best Local Similarity 100.0%;  Pred. No. 2.30e+00;
  Matches     7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    605 FTRQRQP 611
          |||||||
Qy      3 FTRQRQP 9


RESULT    9
ID   A32576     STANDARD;     PRT;   648 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: a32576  check: 9124  from: 1  to: 648.
XX
CC     TOIG of: a32576  check: 9124  from: 1  to: 648
CC
CC    >P1;A32576
CC    beta-glucuronidase (EC 3.2.1.31) B - mouse
CC    C;Species: Mus musculus (house mouse)
CC    C;Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 31-Oct-19
  97
CC    C;Accession: A32576
CC    R;Wawrzyniak, C.J.; Gallagher, P.M.; D'Amore, M.A.; Carter, J.E.; Lund, S
  .D.; Rinchik, E.M.; Ganschow, R.E.
CC    Mol. Cell. Biol. 9, 4074-4078, 1989
CC    A;Title: DNA determinants of structural and regulatory variation within t
  he murine beta-glucuronidase gene complex.
CC    A;Reference number: A32576; MUID:89384641
CC    A;Accession: A32576
```

```
CC    A;Status: preliminary
CC    A;Molecule type: mRNA
CC    A;Residues: 1-648 <WAW>
CC    A;Cross-references: GB:M28540; GB:M27816; NID:g193718; PID:g193719
CC    C;Superfamily: beta-glucuronidase
CC    C;Keywords: glycosidase; hydrolase
SQ   SEQUENCE   648 AA; 74195 MW; 2298575 CN;

  Query Match           60.2%;  Score 56;  DB 2;  Length 648;
  Best Local Similarity 100.0%;  Pred. No. 2.30e+00;
  Matches     7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    605 FTRQRQP 611
          |||||||
Qy      3 FTRQRQP 9


RESULT   10
ID   B32576     STANDARD;     PRT;   648 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: b32576  check: 8794  from: 1  to: 648.
XX
CC     TOIG of: b32576  check: 8794  from: 1  to: 648
CC
CC    >P1;B32576
CC    beta-glucuronidase (EC 3.2.1.31) H - mouse
CC    C;Species: Mus musculus (house mouse)
CC    C;Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 08-Sep-19
  97
CC    C;Accession: B32576
CC    R;Wawrzyniak, C.J.; Gallagher, P.M.; D'Amore, M.A.; Carter, J.E.; Lund, S
  .D.; Rinchik, E.M.; Ganschow, R.E.
CC    Mol. Cell. Biol. 9, 4074-4078, 1989
CC    A;Title: DNA determinants of structural and regulatory variation within t
  he murine beta-glucuronidase gene complex.
CC    A;Reference number: A32576; MUID:89384641
CC    A;Accession: B32576
CC    A;Status: preliminary
CC    A;Molecule type: mRNA
CC    A;Residues: 1-648 <WAW>
CC    A;Cross-references: GB:M28541; NID:g193720; PID:g193721; GB:M27816
CC    C;Superfamily: beta-glucuronidase
CC    C;Keywords: glycosidase; hydrolase
SQ   SEQUENCE   648 AA; 74207 MW; 2297966 CN;

  Query Match           60.2%;  Score 56;  DB 2;  Length 648;
  Best Local Similarity 100.0%;  Pred. No. 2.30e+00;
  Matches     7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    605 FTRQRQP 611
          |||||||
Qy      3 FTRQRQP 9


RESULT   11
ID   S54380     STANDARD;     PRT;   96 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s54380  check: 8538  from: 1  to: 96.
XX
CC     TOIG of: s54380  check: 8538  from: 1  to: 96
CC
CC    >P1;S54380
CC    vpr protein - human immunodeficiency virus type 1
CC    C;Species: human immunodeficiency virus type 1, HIV-1
CC    C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 08-Sep-19
  97
CC    C;Accession: S54380
CC    R;Theodore, T.; Buckler-White, A.J.
CC    submitted to the EMBL Data Library, July 1989
CC    A;Reference number: S54377
CC    A;Accession: S54380
CC    A;Status: preliminary
CC    A;Molecule type: genomic RNA
CC    A;Residues: 1-96 <THE>
CC    A;Cross-references: EMBL:M22639; NID:g329377; PID:g329383
CC    C;Superfamily: AIDS vpr protein
SQ   SEQUENCE   96 AA; 11380 MW; 42519 CN;

  Query Match           59.1%;  Score 55;  DB 2;  Length 96;
  Best Local Similarity 57.1%;  Pred. No. 3.49e+00;
  Matches     8;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGSS 94
          | :||||:: :|||
Qy      1 INFTRQRQPSEGSS 14


RESULT   12
ID   S15047     STANDARD;     PRT;   1703 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S16820.
XX
CC    A;Accession: S16820
CC    A;Molecule type: DNA
CC    A;Residues: 1-1703 <YOS>
CC    A;Cross-references: EMBL:X57837; NID:g4499; PID:g4500
CC    R;Cheret, G.; Sor, F.
CC    submitted to the Protein Sequence Database, July 1996
CC    A;Reference number: S67169
CC    A;Accession: S67192
CC    A;Molecule type: DNA
CC    A;Residues: 1-1703 <CHE>
CC    A;Cross-references: EMBL:Z75198; NID:g1420643; PID:e252424; PID:g1420644;
  MIPS:YOR290c
CC    A;Experimental source: strain S288C
CC    R;Cziepluch, C.; Jauniaux, J.C.; Kordes, E.; Poirey, R.; Pujol, A.; Tobia
  sch, E.
CC    submitted to the Protein Sequence Database, July 1996
CC    A;Reference number: S67194
CC    A;Accession: S67194
CC    A;Molecule type: DNA
CC    A;Residues: 1-308 <CZI>
CC    A;Cross-references: EMBL:Z75198; MIPS:YOR290c
CC    A;Experimental source: strain S288C
CC    R;Cheret, G.; Bernardi, A.; Sor, F.
CC    Yeast 12, 1059-1064, 1996
CC    A;Title: DNA sequence analysis of the VPH1-SNF2 region on chromosome XV o
  f Saccharomyces cerevisiae.
CC    A;Reference number: S72039
CC    A;Accession: S72058
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-1703 <CHW>
CC    A;Cross-references: EMBL:X89633
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, J
  une 1995
CC    C;Genetics:
CC    A;Gene: SGD:SNF2; GAM1
CC    A;Cross-references: MIPS:YOR290c; SGD:S0005816
CC    A;Map position: 15R
CC    C;Superfamily: unassigned bromodomain proteins; bromodomain homology
CC    C;Keywords: nucleus; transcription regulation
CC    F;1576-1631/Domain: bromodomain homology <BRO>
SQ   SEQUENCE   1703 AA; 194050 MW; 14055514 CN;
```

```
  Query Match           59.1%;  Score 55;  DB 2;  Length 1703;
  Best Local Similarity  50.0%;  Pred. No. 3.49e+00;
  Matches     5;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db    299 EFARRRQPTD 308
          :|:|:|||::
Qy      2 NFTRQRQPSE 11


RESULT   13
ID   S46450     STANDARD;     PRT;   186 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   This is a DE line.
XX
SQ   SEQUENCE   186 AA; 20952 MW; 191295 CN;

  Query Match           58.1%;  Score 54;  DB 2;  Length 186;
  Best Local Similarity  63.6%;  Pred. No. 5.27e+00;
  Matches     7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    159 FTRQHQPQDPS 169
          ||||:|| : |
Qy      3 FTRQRQPSEGS 13


RESULT   14
ID   S60396     STANDARD;     PRT;   355 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Cross-references: EMBL:U23084; NID:g1050853; PID:g1050856.
XX
CC    A;Cross-references: EMBL:U23084; NID:g1050853; PID:g1050856
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1995
CC    R;Maurer, C.T.C.; Urbanus, J.H.M.; Planta, R.J.
CC    submitted to the Protein Sequence Database, April 1996
CC    A;Reference number: S63266
CC    A;Accession: S63280
CC    A;Molecule type: DNA
CC    A;Residues: 1-355 <MAW>
CC    A;Cross-references: EMBL:Z71580; NID:g1302399; PID:e239733; PID:g1302400;
   MIPS:YNL304w
CC    A;Experimental source: strain S288C
CC    C;Genetics:
CC    A;Map position: 14L
SQ   SEQUENCE   355 AA; 40678 MW; 624416 CN;

  Query Match           58.1%;  Score 54;  DB 2;  Length 355;
  Best Local Similarity  87.5%;  Pred. No. 5.27e+00;
  Matches     7;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    321 FNFTRQRQ 328
          :|||||||
Qy      1 INFTRQRQ 8


RESULT   15
ID   A69158     STANDARD;     PRT;   373 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: A69158.
XX
CC    A;Accession: A69158
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-373 <MTH>
CC    A;Cross-references: GB:AE000666
CC    A;Experimental source: strain Delta H
CC    C;Genetics:
CC    A;Gene: MTH444
SQ   SEQUENCE   373 AA; 42263 MW; 688580 CN;

  Query Match           58.1%;  Score 54;  DB 2;  Length 373;
  Best Local Similarity  60.0%;  Pred. No. 5.27e+00;
  Matches     6;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    277 IKFTRDRDPA 286
          |:|||:|:|:
Qy      1 INFTRQRQPS 10


RESULT   16
ID   B53293     STANDARD;     PRT;   167 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: B53293.
XX
CC    A;Accession: B53293
CC    A;Status: preliminary
CC    A;Molecule type: DNA
CC    A;Residues: 1-167 <DOL>
CC    A;Cross-references: GB:L19521
CC    C;Genetics:
CC    A;Start codon: GTT
SQ   SEQUENCE   167 AA; 18772 MW; 146186 CN;

  Query Match           57.0%;  Score 53;  DB 2;  Length 167;
  Best Local Similarity  50.0%;  Pred. No. 7.93e+00;
  Matches     5;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db     74 VNYARQRQAG 83
          :|::||||::
Qy      1 INFTRQRQPS 10


RESULT   17
ID   THE INITIATION OF TRANSCRIPTION; THE COMPLEX ALSO REGULATES TRP REPRESSOR
  BIOSYNTHESIS BY BINDING TO ITS REGULATORY REGION     STANDARD;     PRT;   205
   AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: The complete genome sequence of Escherichia coli K-12.
XX
CC    A;Title: The complete genome sequence of Escherichia coli K-12.
CC    A;Reference number: A64720; MUID:97426617
CC    A;Accession: H65254
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-108 <BLAT>
CC    A;Cross-references: GB:AE000509; GB:U00096; NID:g2367383; PID:g1790854; U
  WGP:b4393
CC    A;Experimental source: strain K-12, substrain MG1655
CC    C;Genetics:
CC    A;Gene: trpR
CC    A;Map position: 100 min
CC    C;Function:
CC    A;Description: an aporepressor; when complexed with L-tryptophan it binds
```

```
   the operator region of the trp operon and prevents
SQ   SEQUENCE   205 AA; 23624 MW; 206628 CN;

  Query Match           57.0%;  Score 53;  DB 1;  Length 188;
  Best Local Similarity  53.8%;  Pred. No. 7.93e+00;
  Matches     7;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db     10 ILYTRPREPRESS 22
          | :|| |:| |:|
Qy      1 INFTRQRQPSEGS 13


RESULT   18
ID   F69442     STANDARD;     PRT;   271 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: F69442.
XX
CC    A;Accession: F69442
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-271 <KLE>
CC    A;Cross-references: GB:AE000782; TIGR:AF1543
SQ   SEQUENCE   271 AA; 31009 MW; 390830 CN;

  Query Match           57.0%;  Score 53;  DB 2;  Length 271;
  Best Local Similarity  41.7%;  Pred. No. 7.93e+00;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db     11 DFVRHRQNEDGG 22
          :| |:||  :|:
Qy      2 NFTRQRQPSEGS 13


RESULT   19
ID   A30347     STANDARD;     PRT;   638 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: A30347.
XX
CC    A;Accession: A30347
CC    A;Status: preliminary
CC    A;Molecule type: DNA
CC    A;Residues: 1-638 <GRA>
CC    C;Keywords: exotoxin
SQ   SEQUENCE   638 AA; 69308 MW; 1915001 CN;

  Query Match           57.0%;  Score 53;  DB 2;  Length 638;
  Best Local Similarity  85.7%;  Pred. No. 7.93e+00;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    297 FTRHRQP 303
          |||:|||
Qy      3 FTRQRQP 9


RESULT   20
ID   S24721     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S24721.
XX
CC    A;Accession: S24721
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-317 <ZHO>
CC    A;Cross-references: EMBL:X64851; NID:g44848; PID:g44849
CC    A;Experimental source: strain FA19
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1992
CC    C;Genetics:
CC    A;Gene: fbp
CC    C;Superfamily: sfuA protein
CC    C;Keywords: iron binding
SQ   SEQUENCE   317 AA; 34244 MW; 507109 CN;

  Query Match           53.8%;  Score 50;  DB 2;  Length 317;
  Best Local Similarity  50.0%;  Pred. No. 2.63e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    227 LNFVRHRDPG 236
          :|| |:|:|:
Qy      1 INFTRQRQPS 10


RESULT   21
ID   S24747     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S24747.
XX
CC    A;Accession: S24747
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-317 <ZHO>
CC    A;Cross-references: EMBL:X64852; NID:g45244; PID:g45245
CC    A;Experimental source: strain S3446
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1992
CC    C;Genetics:
CC    A;Gene: fbp
CC    C;Superfamily: sfuA protein
CC    C;Keywords: iron binding
SQ   SEQUENCE   317 AA; 34246 MW; 505589 CN;

  Query Match           53.8%;  Score 50;  DB 2;  Length 317;
  Best Local Similarity  50.0%;  Pred. No. 2.63e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    227 LNFVRHRDPG 236
          :|| |:|:|:
Qy      1 INFTRQRQPS 10


RESULT   22
ID   S24732     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S24732.
XX
CC    A;Accession: S24732
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-317 <ZHO>
CC    A;Cross-references: EMBL:X64853; NID:g45051; PID:g45052
CC    A;Experimental source: strain HF130
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1992
CC    A;Accession: S24725
CC    A;Status: nucleic acid sequence not shown; translation not shown
```

```
CC    A;Molecule type: DNA
CC    A;Residues: 1-317 <ZH2>
CC    A;Cross-references: EMBL:X64858; NID:g45013; PID:g45014
CC    A;Experimental source: strain 44/46
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1992
CC    A;Accession: S24729
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-317 <ZH3>
CC    A;Cross-references: EMBL:X64857; NID:g45045; PID:g45046
CC    A;Experimental source: strain HF116
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1992
CC    A;Accession: S24735
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-317 <ZH4>
CC    A;Cross-references: EMBL:X64856; NID:g45057; PID:g45058
CC    A;Experimental source: strain HF46
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1992
CC    A;Accession: S24738
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-317 <ZH5>
CC    A;Cross-references: EMBL:X64859; NID:g45101; PID:g45102
CC    A;Experimental source: strain M470
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1992
CC    A;Accession: S24740
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-317 <ZH6>
CC    A;Cross-references: EMBL:X64855; NID:g45157; PID:g45158
CC    A;Experimental source: strain N94II
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1992
CC    A;Accession: S24744
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-317 <ZH7>
CC    A;Cross-references: EMBL:X64854; NID:g45173; PID:g45174
CC    A;Experimental source: strain P63
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  arch 1992
CC    C;Genetics:
CC    A;Gene: fbp
CC    C;Superfamily: sfuA protein
CC    C;Keywords: iron binding
SQ   SEQUENCE   317 AA; 34232 MW; 507979 CN;

  Query Match           53.8%;  Score 50;  DB 2;  Length 317;
  Best Local Similarity  50.0%;  Pred. No. 2.63e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    227 LNFVRHRDPG 236
          :|| |:|:|:
Qy      1 INFTRQRQPS 10


RESULT   23
ID   S10256     STANDARD;     PRT;   330 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S10256.
XX
CC    A;Accession: S10256
CC    A;Molecule type: DNA
CC    A;Residues: 1-330 <BER>
CC    A;Cross-references: EMBL:X51901; NID:g44857; PID:g44858
CC    R;Morse, S.A.; Mietzner, T.A.; Bolen, G.; LeFaou, A.; Schoolnik, G.
CC    Antonie Van Leeuwenhoek 53, 465-469, 1987
CC    A;Title: Characterization of the major iron-regulated protein of Neisseri
  a gonorrhoeae and Neissereria meningitidis.
CC    A;Reference number: A60816
CC    A;Accession: A60816
CC    A;Molecule type: protein
CC    A;Residues: 23-69 <MOR>
CC    R;Mietzner, T.A.; Bolan, G.; Schoolnik, G.K.; Morse, S.A.
CC    J. Exp. Med. 165, 1041-1057, 1987
CC    A;Title: Purification and characterization of the major iron-regulated pr
  otein expressed by pathogenic Neisseriae.
CC    A;Reference number: A46544
CC    A;Accession: A46544
CC    A;Status: preliminary
CC    A;Molecule type: protein
CC    A;Residues: 23-69 <MIE>
CC    C;Genetics:
CC    A;Gene: Fbp
CC    C;Superfamily: sfuA protein
CC    C;Keywords: iron binding
CC    F;1-22/Domain: signal sequence #status experimental <SIG>
CC    F;23-330/Product: iron-binding protein #status experimental <MAT>
SQ   SEQUENCE   330 AA; 35769 MW; 546828 CN;

  Query Match           53.8%;  Score 50;  DB 2;  Length 330;
  Best Local Similarity  50.0%;  Pred. No. 2.63e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    235 LNFVRHRDPG 244
          :|| |:|:|:
Qy      1 INFTRQRQPS 10


RESULT   24
ID   B60816     STANDARD;     PRT;   330 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: b60816  check: 1589  from: 1  to: 330.
XX
CC     TOIG of: b60816  check: 1589  from: 1  to: 330
CC
CC    >P1;B60816
CC    major iron-regulated protein - Neisseria meningitidis
CC    N;Alternate names: iron-binding protein; MIRP
CC    C;Species: Neisseria meningitidis
CC    C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Sep-19
  97
CC    C;Accession: S10978; B60816
CC    R;Berish, S.A.; Kapczynski, D.R.; Morse, S.A.
CC    Nucleic Acids Res. 18, 4596, 1990
CC    A;Title: Nucleotide sequence of the Fbp gene from Neisseria meningitidis.
CC    A;Reference number: S10978; MUID:90356404
CC    A;Accession: S10978
CC    A;Molecule type: DNA
CC    A;Residues: 1-330 <BER>
CC    A;Cross-references: EMBL:X53467; NID:g45037; PID:g45038
CC    R;Morse, S.A.; Mietzner, T.A.; Bolen, G.; LeFaou, A.; Schoolnik, G.
CC    Antonie Van Leeuwenhoek 53, 465-469, 1987
CC    A;Title: Characterization of the major iron-regulated protein of Neisseri
  a gonorrhoeae and Neissereria meningitidis.
CC    A;Reference number: A60816
CC    A;Accession: B60816
CC    A;Molecule type: protein
CC    A;Residues: 23-69 <MOR>
CC    C;Genetics:
CC    A;Gene: Fbp
CC    C;Superfamily: sfuA protein
CC    C;Keywords: iron binding
```

```
CC    F;23-330/Product: major iron-regulated protein #status experimental <MAT>
SQ   SEQUENCE   330 AA; 35741 MW; 544204 CN;

  Query Match           53.8%;  Score 50;  DB 2;  Length 330;
  Best Local Similarity  50.0%;  Pred. No. 2.63e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    235 LNFVRHRDPG 244
          :|| |:|:|:
Qy      1 INFTRQRQPS 10


RESULT   25
ID   A48585     STANDARD;     PRT;   410 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: a48585  check: 4186  from: 1  to: 410.
XX
CC     TOIG of: a48585  check: 4186  from: 1  to: 410
CC
CC    >P1;A48585
CC    transcription factor DP-1 - human
CC    C;Species: Homo sapiens (man)
CC    C;Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 10-Sep-19
  97
CC    C;Accession: A48585
CC    R;Helin, K.; Wu, C.L.; Fattaey, A.R.; Lees, J.A.; Dynlacht, B.D.; Ngwu, C
  .; Harlow, E.
CC    Genes Dev. 7, 1850-1861, 1993
CC    A;Title: Heterodimerization of the transcription factors E2F-1 and DP-1 l
  eads to cooperative trans-activation.
CC    A;Reference number: A48585
CC    A;Accession: A48585
CC    A;Status: preliminary; nucleic acid sequence not shown
CC    A;Molecule type: mRNA
CC    A;Residues: 1-410 <HEL>
CC    A;Cross-references: GB:L23959; NID:g414316; PID:g414317
CC    C;Keywords: DNA binding; transcription factor
SQ   SEQUENCE   410 AA; 45070 MW; 833309 CN;

  Query Match           53.8%;  Score 50;  DB 2;  Length 410;
  Best Local Similarity  46.2%;  Pred. No. 2.63e+01;
  Matches     6;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     87 HFASQNQPSDSSP 99
          :|: | |||::|:
Qy      2 NFTRQRQPSEGSS 14


RESULT   26
ID   SURFCA     STANDARD;     PRT;   793 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: surfca  check: 4832  from: 1  to: 793.
XX
CC     TOIG of: surfca  check: 4832  from: 1  to: 793
CC
CC    >P1;SURFCA
CC    ATP-dependent Clp proteinase (EC 3.4.21.-) chain A homolog - Rhodopseudom
  onas blastica
CC    C;Species: Rhodopseudomonas blastica
CC    C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 28-Feb-19
  97
CC    C;Accession: S04667
CC    R;Tybulewicz, V.L.J.; Falk, G.; Walker, J.E.
CC    J. Mol. Biol. 179, 185-214, 1984
CC    A;Title: Rhodopseudomonas blastica atp operon. Nucleotide sequence and tr
  anscription.
CC    A;Reference number: S04666; MUID:85058188
CC    A;Accession: S04667
CC    A;Molecule type: DNA
CC    A;Residues: 1-793 <TYB>
CC    C;Superfamily: ATP-dependent Clp proteinase chain A
CC    C;Keywords: ATP; hydrolase; nucleotide binding; P-loop; serine proteinase
CC    F;244-252/Region: nucleotide-binding motif A (P-loop)
CC    F;311-315/Region: nucleotide-binding motif B
CC    F;525-533/Region: nucleotide-binding motif A (P-loop)
CC    F;590-594/Region: nucleotide-binding motif B
SQ   SEQUENCE   793 AA; 86890 MW; 3017877 CN;

  Query Match           53.8%;  Score 50;  DB 1;  Length 793;
  Best Local Similarity  35.7%;  Pred. No. 2.63e+01;
  Matches     5;  Conservative    5;  Mismatches   4;  Indels   0;  Gaps   0;

Db    647 IGFGRERRTGEDTA 660
          | | |:|: :| ::
Qy      1 INFTRQRQPSEGSS 14


RESULT   27
ID   A41259     STANDARD;     PRT;   889 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: a41259  check: 4909  from: 1  to: 889.
XX
CC     TOIG of: a41259  check: 4909  from: 1  to: 889
CC
CC    >P1;A41259
CC    potassium transport protein TRK2, low-affinity - yeast (Saccharomyces cer
  evisiae)
CC    N;Alternate names: protein YKR050w
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 05-Dec-19
  97
CC    C;Accession: A41259; S38124; S17489
CC    R;Ko, C.H.; Gaber, R.F.
CC    Mol. Cell. Biol. 11, 4266-4273, 1991
CC    A;Title: TRK1 and TRK2 encode structurally related K(+) transporters in S
  accharomyces cerevisiae.
CC    A;Reference number: A41259; MUID:91304419
CC    A;Accession: A41259
CC    A;Molecule type: DNA
CC    A;Residues: 1-889 <KOC>
CC    A;Cross-references: GB:M65215; NID:g173032; PID:g173033
CC    R;Vissers, S.; Urrestarazu, L.A.; Jauniaux, J.C.
CC    submitted to the Protein Sequence Database, March 1994
CC    A;Reference number: S38118
CC    A;Accession: S38124
CC    A;Molecule type: DNA
CC    A;Residues: 1-889 <VIS>
CC    A;Cross-references: EMBL:Z28275; NID:g486502; PID:g486503; MIPS:YKR050w
CC    A;Experimental source: strain S288C
CC    C;Genetics:
CC    A;Gene: SGD:TRK2
CC    A;Cross-references: SGD:S0001758
CC    A;Map position: 11R
CC    C;Superfamily: potassium transport protein TRK1
CC    C;Keywords: ion transport; potassium transport; transmembrane protein
CC    F;53-73/Domain: transmembrane #status predicted <TM01>
CC    F;80-101/Domain: transmembrane #status predicted <TM02>
CC    F;109-131/Domain: transmembrane #status predicted <TM03>
CC    F;465-487/Domain: transmembrane #status predicted <TM04>
CC    F;500-521/Domain: transmembrane #status predicted <TM05>
CC    F;525-545/Domain: transmembrane #status predicted <TM06>
CC    F;549-569/Domain: transmembrane #status predicted <TM07>
CC    F;585-605/Domain: transmembrane #status predicted <TM08>
CC    F;609-630/Domain: transmembrane #status predicted <TM09>
```

```
CC    F;658-662/Domain: transmembrane #status predicted <TM10>
CC    F;744-764/Domain: transmembrane #status predicted <TM11>
CC    F;777-797/Domain: transmembrane #status predicted <TM12>
SQ   SEQUENCE   889 AA; 101086 MW; 4294395 CN;

  Query Match           53.8%;  Score 50;  DB 2;  Length 889;
  Best Local Similarity  60.0%;  Pred. No. 2.63e+01;
  Matches     6;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    233 NFSSKRQPSD 242
          ||:  |||||:
Qy      2 NFTRQRQPSE 11


RESULT   28
ID   DING EGG-LAYING HORMONE, THAT CONTROL THE PHYSIOLOGICAL PROCESSES OF EGG-L
  AYING.     STANDARD;     PRT;   132 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: D26306.
XX
CC    A;Accession: D26306
CC    A;Molecule type: protein
CC    A;Residues: 117-152,156-173 <NA2>
CC    C;Comment: Atrial gland peptides A and B initiate egg-laying by stimulati
  ng bag cell neurons to release other peptides, incl
SQ   SEQUENCE   132 AA; 15099 MW; 109782 CN;

  Query Match           52.7%;  Score 49;  DB 1;  Length 102;
  Best Local Similarity  41.7%;  Pred. No. 3.89e+01;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     72 ILYAREREPRES 83
          | ::|:|:| |:
Qy      1 INFTRQRQPSEG 12


RESULT   29
ID   DING EGG-LAYING HORMONE, THAT CONTROL THE PHYSIOLOGICAL PROCESSES OF EGG-L
  AYING.     STANDARD;     PRT;   132 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: B26306.
XX
CC    A;Accession: B26306
CC    A;Molecule type: protein
CC    A;Residues: 117-152;156-173 <NA2>
CC    C;Comment: Atrial gland peptides A and B initiate egg-laying by stimulati
  ng bag cell neurons to release other peptides, incl
SQ   SEQUENCE   132 AA; 15099 MW; 109782 CN;

  Query Match           52.7%;  Score 49;  DB 1;  Length 102;
  Best Local Similarity  41.7%;  Pred. No. 3.89e+01;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     72 ILYAREREPRES 83
          | ::|:|:| |:
Qy      1 INFTRQRQPSEG 12


RESULT   30
ID   NRYY     STANDARD;     PRT;   128 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: A00824.
XX
CC    A;Accession: A00824
CC    A;Molecule type: protein
CC    A;Residues: 1-128 <BEI>
CC    C;Comment: The capybara is a rodent belonging to the same superfamily as
  the guinea pig.
CC    C;Superfamily: pancreatic ribonuclease
CC    C;Keywords: hydrolase; nucleic acid digestion; pancreas
CC    F;12,41,119/Active site: His, Lys, His #status predicted
CC    F;26-84,40-95,58-110,65-72/Disulfide bonds: #status predicted
SQ   SEQUENCE   128 AA; 14345 MW; 92128 CN;

  Query Match           52.7%;  Score 49;  DB 1;  Length 128;
  Best Local Similarity  57.1%;  Pred. No. 3.89e+01;
  Matches     8;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db      6 MKFQRQHVDSEGSS 19
          ::| ||:  |||||
Qy      1 INFTRQRQPSEGSS 14


RESULT   31
ID   S03523     STANDARD;     PRT;   130 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s03523  check: 7915  from: 1  to: 130.
XX
CC     TOIG of: s03523  check: 7915  from: 1  to: 130
CC
CC    >F1;S03523
CC    T-cell receptor alpha chain precursor V-J region (HD-Mar) - human (fragme
  nt)
CC    C;Species: Homo sapiens (man)
CC    C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Sep-19
  97
CC    C;Accession: S03523
CC    R;Luria, S.; Gross, G.; Horowitz, M.; Givol, D.
CC    EMBO J. 6, 3307-3312, 1987
CC    A;Title: Promoter and enhancer elements in the rearranged alpha chain gen
  e of the human T cell receptor.
CC    A;Reference number: S03523; MUID:88111516
CC    A;Accession: S03523
CC    A;Molecule type: DNA
CC    A;Residues: 1-130 <LUR>
CC    A;Cross-references: EMBL:X06192; NID:g36908; PID:g296677
CC    A;Note: this sequence was determined from the differentiated gene
CC    C;Genetics:
CC    A;Introns: 15/1
CC    C;Superfamily: immunoglobulin V region; immunoglobulin homology
CC    C;Keywords: T-cell receptor
CC    F;1-18/Domain: signal sequence #status predicted <SIG>
CC    F;19-110/Product: T-cell receptor alpha chain V region HD-Mar #status pre
  dicted <MAT>
CC    F;114-130/Domain: J region (J-alpha-3) #status experimental <JRE>
SQ   SEQUENCE   130 AA; 14687 MW; 94561 CN;

  Query Match           52.7%;  Score 49;  DB 2;  Length 130;
  Best Local Similarity  41.7%;  Pred. No. 3.89e+01;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     94 FIRDSQPSDSAT 105
          | |: |||::::
Qy      3 FTRQRQPSEGSS 14


RESULT   32
ID   A48912     STANDARD;     PRT;   237 AA.
XX
AC   xxxxxx
```

```
XX
DT   01-JAN-1900
XX
DE   TOIG of: a48912  check: 632  from: 1  to: 237.
XX
CC     TOIG of: a48912  check: 632  from: 1  to: 237
CC
CC    >P1;A48912
CC    leucine zipper protein Nrl, retina-specific - mouse
CC    C;Species: Mus musculus (house mouse)
CC    C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Sep-19
  97
CC    C;Accession: A48912
CC    R;Farjo, Q.; Jackson, A.U.; Xu, J.; Gryzenia, M.; Skolnick, C.; Agarwal,
  N.; Swaroop, A.
CC    Genomics 18, 216-222, 1993
CC    A;Title: Molecular characterization of the murine neural retina leucine z
  ipper gene, Nrl.
CC    A;Reference number: A48912
CC    A;Accession: A48912
CC    A;Status: preliminary
CC    A;Molecule type: mRNA
CC    A;Residues: 1-237 <FAR>
CC    A;Cross-references: GB:L14935; NID:g388916; PID:g388917
CC    C;Genetics:
CC    A;Gene: Nrl
CC    C;Superfamily: maf transforming protein; maf homology
CC    C;Keywords: DNA binding; leucine zipper; retina
CC    F;133-222/Domain: maf homology <MAF>
SQ   SEQUENCE   237 AA; 26083 MW; 238917 CN;

  Query Match           52.7%;  Score 49;  DB 2;  Length 237;
  Best Local Similarity  50.0%;  Pred. No. 3.89e+01;
  Matches     7;  Conservative    3;  Mismatches   4;  Indels   0;  Gaps   0;

Db     19 MKFEIKREPSEGRS 32
          ::|   |:|||| |
Qy      1 INFTRQRQPSEGSS 14


RESULT   33
ID   I68889     STANDARD;     PRT;   314 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: PH1294.
XX
CC    A;Accession: PH1294
CC    A;Molecule type: DNA
CC    A;Residues: 168-176 <TRA>
CC    C;Genetics:
CC    A;Gene: GDB:MAGEA2; MAGE2
CC    A;Cross-references: GDB:273684
CC    A;Map position: Xq28-Xq28
SQ   SEQUENCE   314 AA; 35054 MW; 525196 CN;

  Query Match           52.7%;  Score 49;  DB 2;  Length 314;
  Best Local Similarity  64.3%;  Pred. No. 3.89e+01;
  Matches     9;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db     75 INYTLWRQSDEGSS 88
          ||:|  ||: ||||
Qy      1 INFTRQRQPSEGSS 14


RESULT   34
ID   S74937     STANDARD;     PRT;   343 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S74937.
XX
CC    A;Accession: S74937
CC    A;Status: preliminary
CC    A;Molecule type: DNA
CC    A;Residues: 1-343 <KAN>
CC    A;Cross-references: EMBL:D90902; NID:g1652027; PID:d1017710; PID:g1652052
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, J
  une 1996
SQ   SEQUENCE   343 AA; 37956 MW; 575559 CN;

  Query Match           52.7%;  Score 49;  DB 2;  Length 343;
  Best Local Similarity  87.5%;  Pred. No. 3.89e+01;
  Matches     7;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Db    315 RQRQPSTG 322
          |||||| |
Qy      5 RQRQPSEG 12


RESULT   35
ID   A28658     STANDARD;     PRT;   349 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: A28658.
XX
CC    A;Accession: A28658
CC    A;Molecule type: DNA
CC    A;Residues: 1-349 <STA>
CC    A;Cross-references: GB:J03196; NID:g149174; PID:g149175
CC    C;Keywords: hydrolase
SQ   SEQUENCE   349 AA; 37801 MW; 607988 CN;

  Query Match           52.7%;  Score 49;  DB 2;  Length 349;
  Best Local Similarity  50.0%;  Pred. No. 3.89e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    298 VSINRQRQPA 307
          ::: |||||:
Qy      1 INFTRQRQPS 10


Search completed: Tue Apr  7 08:38:47 1998
Job time : 15 secs.
```

****

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:37:29 1998; MasPar time 2.36 Seconds
148.813 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-2
Description: (1-14) from 5541104.pep
Perfect Score: 93
Sequence: 1 INFTRQRQPSEGSS 14

Scoring table: PAM 150
Gap 15

Searched: 69112 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 25.257; Variance 27.960; scale 0.903

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	93	100.0	309	1	MAG1_HUMAN	MELANOMA-ASSOCIATED AN	4.59e-11
2	58	62.4	172	1	RL10_BRUAB	50S RIBOSOMAL PROTEIN	1.04e-01
3	57	61.3	317	1	MAG4_HUMAN	MELANOMA-ASSOCIATED AN	1.77e-01
4	57	61.3	651	1	BGLR_HUMAN	BETA-GLUCURONIDASE PRE	1.77e-01
5	56	60.2	648	1	BGLR_MOUSE	BETA-GLUCURONIDASE PRE	3.00e-01
6	55	59.1	96	1	VPR_HV1Z2	VPR PROTEIN (R ORF PRO	5.07e-01
7	55	59.1	96	1	VPR_HV1MA	VPR PROTEIN (R ORF PRO	5.07e-01
8	55	59.1	1703	1	SNF2_YEAST	TRANSCRIPTION REGULATO	5.07e-01
9	54	58.1	355	1	YN44_YEAST	HYPOTHETICAL 40.7 KD G	8.49e-01
10	53	57.0	96	1	VPR_HV1OY	VPR PROTEIN (R ORF PRO	1.41e+00
11	53	57.0	96	1	VPR_HV1JR	VPR PROTEIN (R ORF PRO	1.41e+00
12	53	57.0	96	1	VPR_HV1RH	VPR PROTEIN (R ORF PRO	1.41e+00
13	53	57.0	511	1	ACH5_CAEEL	ACETYLCHOLINE RECEPTOR	1.41e+00
14	53	57.0	638	1	TOXA_PSEAE	EXOTOXIN A PRECURSOR (	1.41e+00
15	51	54.8	331	1	SRA9_CAEEL	SRA-9 PROTEIN.	3.83e+00
16	50	53.8	96	1	VPR_HV1EL	VPR PROTEIN (R ORF PRO	6.23e+00
17	50	53.8	330	1	FBP_NEIGO	MAJOR FERRIC IRON BIND	6.23e+00
18	50	53.8	330	1	FBP_NEIME	MAJOR FERRIC IRON BIND	6.23e+00
19	50	53.8	793	1	CLPA_RHOBL	CLPA HOMOLOG PROTEIN.	6.23e+00
20	50	53.8	889	1	TRK2_YEAST	POTASSIUM TRANSPORT PR	6.23e+00
21	49	52.7	96	1	VPR_HV1N5	VPR PROTEIN (R ORF PRO	1.01e+01
22	49	52.7	128	1	RNP_HYDHY	RIBONUCLEASE PANCREATI	1.01e+01
23	49	52.7	155	1	RS7_MYCSM	30S RIBOSOMAL PROTEIN	1.01e+01
24	49	52.7	237	1	NRL_MOUSE	NEURAL RETINA-SPECIFIC	1.01e+01
25	49	52.7	314	1	MAG2_HUMAN	MELANOMA-ASSOCIATED AN	1.01e+01
26	49	52.7	349	1	NRLB_KLEPN	NITRILASE, BROMOXYNIL-	1.01e+01
27	49	52.7	494	1	SCRB_STAXY	SUCROSE-6-PHOSPHATE HY	1.01e+01
28	49	52.7	542	1	MET7_NEUCR	CYSTATHIONINE GAMMA-SY	1.01e+01
29	49	52.7	648	1	BGLR_RAT	BETA-GLUCURONIDASE PRE	1.01e+01
30	49	52.7	733	1	ERG7_RAT	LANOSTEROL SYNTHASE (E	1.01e+01
31	48	51.6	94	1	RT19_PETHY	MITOCHONDRIAL RIBOSOMA	1.61e+01
32	48	51.6	96	1	VPR_HV1MN	VPR PROTEIN (R ORF PRO	1.61e+01
33	48	51.6	162	1	PHCA_SYNY3	C-PHYCOCYANIN ALPHA CH	1.61e+01
34	48	51.6	220	1	US02_HSVBS	PROTEIN US2 HOMOLOG.	1.61e+01
35	48	51.6	272	1	MCBC_ECOLI	MICROCIN B17 PROCESSIN	1.61e+01
36	48	51.6	318	1	Y066_MYCTU	HYPOTHETICAL 34.6 KD P	1.61e+01
37	48	51.6	518	1	YNV6_CAEEL	HYPOTHETICAL 59.0 KD P	1.61e+01
38	48	51.6	552	1	HMEN_DROME	SEGMENTATION POLARITY	1.61e+01
39	48	51.6	559	1	3BP2_MOUSE	SH3 DOMAIN-BINDING PRO	1.61e+01
40	48	51.6	584	1	HMEN_DROVI	SEGMENTATION POLARITY	1.61e+01
41	48	51.6	3567	1	ERY2_SACER	ERYTHRONOLIDE SYNTHASE	1.61e+01
42	47	50.5	162	1	HXA9_CAVPO	HOMEOBOX PROTEIN HOX-A	2.56e+01
43	47	50.5	231	1	UCRI_NEUCR	UBIQUINOL-CYTOCHROME C	2.56e+01
44	47	50.5	246	1	YU17_MYCTU	HYPOTHETICAL 25.3 KD P	2.56e+01
45	47	50.5	292	1	GTAB_BACSU	UTP--GLUCOSE-1-PHOSPHA	2.56e+01
46	47	50.5	305	1	YB3_XENLA	B BOX BINDING PROTEIN	2.56e+01
47	47	50.5	317	1	YB1_HUMAN	Y BOX BINDING PROTEIN-	2.56e+01
48	47	50.5	321	1	YB1_CHICK	Y BOX BINDING PROTEIN-	2.56e+01
49	47	50.5	322	1	CBFX_MOUSE	CCAAT-BINDING TRANSCRI	2.56e+01
50	47	50.5	322	1	YB1_MOUSE	Y BOX BINDING PROTEIN-	2.56e+01
51	47	50.5	324	1	CBFX_HUMAN	CCAAT-BINDING TRANSCRI	2.56e+01
52	47	50.5	362	1	SCP3_SCHPO	SCP3 PROTEIN.	2.56e+01
53	47	50.5	459	1	G33_RAT	GENE 33 POLYPEPTIDE.	2.56e+01
54	47	50.5	460	1	V51K_ACLSP	50.8 KD PROTEIN (ORF2)	2.56e+01
55	47	50.5	479	1	OPRJ_PSEAE	OUTER MEMBRANE PROTEIN	2.56e+01
56	47	50.5	506	1	PCCB_BACSU	PUTATIVE PROPIONYL-COA	2.56e+01
57	47	50.5	517	1	T2FA_HUMAN	TRANSCRIPTION INITIATI	2.56e+01
58	47	50.5	630	1	TEC_MOUSE	TYROSINE-PROTEIN KINAS	2.56e+01
59	47	50.5	631	1	TEC_HUMAN	TYROSINE-PROTEIN KINAS	2.56e+01
60	47	50.5	827	1	PLSB_ECOLI	GLYCEROL-3-PHOSPHATE A	2.56e+01
61	47	50.5	1060	1	UAY_EMENI	POSITIVE REGULATOR OF	2.56e+01
62	47	50.5	1144	1	RIR1_HSV23	RIBONUCLEOSIDE-DIPHOSP	2.56e+01
63	47	50.5	1359	1	STH1_YEAST	NUCLEAR PROTEIN STH1/N	2.56e+01
64	46	49.5	152	1	RNP_RAT	RIBONUCLEASE PANCREATI	4.03e+01
65	46	49.5	152	1	ADF2_CAEEL	ACTIN DEPOLYMERIZING F	4.03e+01
66	46	49.5	196	1	TRAP_ECOLI	TRAP PROTEIN.	4.03e+01
67	46	49.5	275	1	RK2_ODOSI	CHLOROPLAST 50S RIBOSO	4.03e+01
68	46	49.5	303	1	YB1_XENLA	Y BOX BINDING PROTEIN-	4.03e+01
69	46	49.5	310	1	YGLA_SYNP2	HYPOTHETICAL 34.1 KD P	4.03e+01
70	46	49.5	320	1	EXON_RHIME	UTP--GLUCOSE-1-PHOSPHA	4.03e+01
71	46	49.5	328	1	YG19_YEAST	HYPOTHETICAL 37.9 KD P	4.03e+01
72	46	49.5	400	1	ARRH_LIMPO	ARRESTIN, LATERAL EYE.	4.03e+01
73	46	49.5	406	1	KDGC_ERWCH	PECTIN DEGRADATION PRO	4.03e+01
74	46	49.5	425	1	K2C1_XENLA	KERATIN, TYPE II CYTOS	4.03e+01
75	46	49.5	428	1	RF4_KLULA	RF4 PROTEIN.	4.03e+01
76	46	49.5	438	1	Y16R_MYCIO	HYPOTHETICAL 51.3 KD P	4.03e+01
77	46	49.5	492	1	CATA_IPOBA	CATALASE (EC 1.11.1.6)	4.03e+01
78	46	49.5	492	1	CAT3_ARATH	CATALASE 3 (EC 1.11.1.	4.03e+01
79	46	49.5	492	1	CAT2_RICCO	CATALASE ISOZYME 2 (EC	4.03e+01
80	46	49.5	504	1	YHV9_YEAST	HYPOTHETICAL 56.3 KD P	4.03e+01
81	46	49.5	832	1	DPO1_THEAQ	DNA POLYMERASE (EC 2.7	4.03e+01
82	46	49.5	1204	1	YAEF_SCHPO	HYPOTHETICAL PROTEIN C	4.03e+01
83	46	49.5	1466	1	NKCR_MOUSE	NK-TUMOR RECOGNITION P	4.03e+01
84	45	48.4	224	1	J1S_HCMVA	HYPOTHETICAL PROTEIN J	6.29e+01
85	45	48.4	231	1	YJIH_ECOLI	HYPOTHETICAL 24.2 KD P	6.29e+01
86	45	48.4	294	1	YGHA_ECOLI	HYPOTHETICAL OXIDOREDU	6.29e+01
87	45	48.4	305	1	SLEB_BACSU	SPORE-CORTEX-LYTIC ENZ	6.29e+01
88	45	48.4	309	1	J1L_HCMVA	HYPOTHETICAL PROTEIN J	6.29e+01
89	45	48.4	331	1	YZ03_HUMAN	HYPOTHETICAL PROTEIN Z	6.29e+01
90	45	48.4	532	1	EST2_RABIT	LIVER CARBOXYLESTERASE	6.29e+01
91	45	48.4	551	1	IL2B_HUMAN	INTERLEUKIN-2 RECEPTOR	6.29e+01
92	45	48.4	589	1	PHBC_ALCEU	POLY-BETA-HYDROXYBUTYR	6.29e+01
93	45	48.4	644	1	YNL5_YEAST	HYPOTHETICAL 74.0 KD P	6.29e+01
94	45	48.4	708	1	MR11_HUMAN	MRE11 HOMOLOG.	6.29e+01
95	45	48.4	732	1	IUTA_ECOLI	FERRIC AEROBACTIN RECE	6.29e+01
96	45	48.4	934	1	C1TC_HUMAN	C-1-TETRAHYDROFOLATE S	6.29e+01

97	45	48.4	1017	1	MCM6_YEAST	MINICHROMOSOME MAINTEN 6.29e+01
98	45	48.4	1226	1	YMA1_CAEEL	PROBABLE INTEGRIN ALPH 6.29e+01
99	45	48.4	3344	1	POLG_PRSVH	GENOME POLYPROTEIN (CO 6.29e+01
100	45	48.4	5217	1	HTS1_COCCA	HC-TOXIN SYNTHETASE (E 6.29e+01

ALIGNMENTS

```
RESULT    1
ID   MAG1_HUMAN     STANDARD;      PRT;   309 AA.
AC   P43355; O00346;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 1 (MAGE-1 ANTIGEN) (ANTIGEN MZ2-E).
GN   MAGEA1 OR MAGE1 OR MAGE1A.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92086861.
RA   VAN DER BRUGGEN P., TRAVERSARI C., CHOMEZ P., LURQUIN C., DE PLAEN E.,
RA   VAN DEN EYNDE B., KNUTH A., BOON T.;
RL   SCIENCE 254:1643-1647(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKIN;
RX   MEDLINE; 94311935.
RA   DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   GLOECKNER G., RUMP A., NORDSIEK G., HINZMANN B., KIOSCHIS P.,
RA   HEISS N., POUSTKA A., BAUER D., DRESCHER B., KNOB A., ROSENTHAL A.;
RL   SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [4]
RP   MUTAGENESIS.
RC   TISSUE=BLOOD;
RX   MEDLINE; 94157413.
RA   GAUGLER B., VAN DEN EYNDE B., VAN DER BRUGGEN P., ROMERO P.,
RA   GAFORIO J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;
RL   J. EXP. MED. 179:921-930(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC       CYTOLYTIC T LYMPHOCYTES.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES. NEVER EXPRESSED IN KIDNEY TUMORS, LEUKEMIAS AND
CC       LYMPHOMAS.
CC   -!- POLYMORPHISM: THE VARIANT AT POSITION 32 LIKELY REPRESENTS A
CC       POLYMORPHISM OF THE MAG-1 GENE.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; M77481; G416115; -.
DR   EMBL; U82672; G2078527; -.
DR   MIM; 300016; -.
KW   ANTIGEN; MULTIGENE FAMILY; POLYMORPHISM; TUMOR ANTIGEN.
FT   VARIANT      32     32       T -> A.
FT   DOMAIN       33     36       POLY-SER.
FT   MUTAGEN     163    163       D->A: ABOLISHES HLA-A1 BINDING.
FT   MUTAGEN     169    169       Y->A: ABOLISHES HLA-A1 BINDING.
FT   CONFLICT     72     72       R -> Q (IN REF. 3).
SQ   SEQUENCE   309 AA;  34342 MW;  E6CB1300 CRC32;

  Query Match          100.0%;  Score 93;  DB 1;  Length 309;
  Best Local Similarity 100.0%;  Pred. No. 4.59e-11;
  Matches    14;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     68 INFTRQRQPSEGSS 81
          ||||||||||||||
Qy      1 INFTRQRQPSEGSS 14
```

```
RESULT    2
ID   RL10_BRUAB     STANDARD;      PRT;   172 AA.
AC   P41107;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   50S RIBOSOMAL PROTEIN L10.
GN   RPLJ.
OS   BRUCELLA ABORTUS.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   UNCERTAIN.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=19;
RX   MEDLINE; 94171071.
RA   OLIVEIRA S.C., ZHU Y., SPLITTER G.A.;
RL   GENE 140:137-138(1994).
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   EMBL; L23505; G387911; -.
DR   PROSITE; PS01109; RIBOSOMAL_L10; 1.
KW   RIBOSOMAL PROTEIN.
SQ   SEQUENCE   172 AA;  18459 MW;  196B1070 CRC32;

  Query Match          62.4%;  Score 58;  DB 1;  Length 172;
  Best Local Similarity 60.0%;  Pred. No. 1.04e-01;
  Matches     6;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db     13 VKFVRQRQPG 22
          ::| |||||:
Qy      1 INFTRQRQPS 10
```

```
RESULT    3
ID   MAG4_HUMAN     STANDARD;      PRT;   317 AA.
AC   P43358;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 4 (MAGE-4 ANTIGEN) (MAGE-X2) (MAGE-41).
GN   MAGEA4 OR MAGE4.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKIN;
RX   MEDLINE; 94311935.
RA   DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95369706.
RA   IMAI Y., SHICHIJO S., YAMADA A., KATAYAMA T., YANO H., ITOH K.;
RL   GENE 160:287-290(1995).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY WITH
CC       MAGE-1.
```

```
DR   EMBL; U10687; G533515; -.
DR   EMBL; U10688; G533517; -.
DR   EMBL; U10340; G499124; -.
DR   EMBL; D32077; G1125018; -.
KW   ANTIGEN; MULTIGENE FAMILY; POLYMORPHISM; TUMOR ANTIGEN.
FT   DOMAIN       41     44       POLY-SER.
FT   VARIANT     173    173       T -> A.
FT   CONFLICT    307    307       E -> Q (IN REF. 2).
SQ   SEQUENCE   317 AA;  34929 MW;  3CE38AF9 CRC32;

  Query Match           61.3%;  Score 57;  DB 1;  Length 317;
  Best Local Similarity  71.4%;  Pred. No. 1.77e-01;
  Matches    10;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db     76 ISFTCWRQPNEGSS 89
          |:||  |||:||||
Qy      1 INFTRQRQPSEGSS 14


RESULT    4
ID   BGLR_HUMAN     STANDARD;      PRT;   651 AA.
AC   P08236;
DT   01-AUG-1988 (REL. 08, CREATED)
DT   01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   BETA-GLUCURONIDASE PRECURSOR (EC 3.2.1.31) (BETA-G1).
GN   GUSB.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87118233.
RA   OSHIMA A., KYLE J.W., MILLER R.D., HOFFMANN J.W., POWELL P.P.,
RA   GRUBB J.H., SLY W.S., TROPAK M., GUISE K.S., GRAVEL R.A.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 84:685-689(1987).
RN   [2]
RP   SEQUENCE OF 1-70 FROM N.A.
RX   MEDLINE; 92009900.
RA   SHIPLEY J.M., MILLER R.D., WU B.M., GRUBB J.H., CHRISTENSEN S.G.,
RA   KYLE J.W., SLY W.S.;
RL   GENOMICS 10:1009-1018(1991).
RN   [3]
RP   SEQUENCE OF 23-32 AND 160-175.
RC   TISSUE=PLACENTA;
RX   MEDLINE; 92162201.
RA   TANAKA J., GASA S., SAKURADA K., MIYAZAKI T., KASAI M., MAKITA A.;
RL   BIOL. CHEM. HOPPE-SEYLER 373:57-62(1992).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX   MEDLINE; 96185449.
RA   JAIN S., DRENDEL W.B., CHEN Z.W., MATHEWS F.S., SLY W.S., GRUBB J.H.;
RL   NAT. STRUCT. BIOL. 3:375-381(1996).
RN   [5]
RP   VARIANT MPS-VII TRP-216.
RX   MEDLINE; 94154730.
RA   VERVOORT R., LISSENS W., LIEBAERS I.;
RL   HUM. MUTAT. 2:443-445(1993).
RN   [6]
RP   VARIANTS MPS-VII VAL-354 AND TRP-611.
RX   MEDLINE; 94154731.
RA   WU B.M., SLY W.S.;
RL   HUM. MUTAT. 2:446-457(1993).
RN   [7]
RP   VARIANTS MPS-VII CYS-382 AND VAL-619.
RX   MEDLINE; 91090114.
RA   TOMATSU S., FUKUDA S., SUKEGAWA K., IKEDO Y., YAMADA S., YAMADA Y.,
RA   SASAKI T., OKAMOTO H., KUWAHARA T., YAMAGUCHI S., KIMAN T.,
RA   SHINTAKU H., ISSHIKI G., ORII T.;
RL   AM. J. HUM. GENET. 48:89-96(1991).
RN   [8]
RP   VARIANT MPS-VII CYS-627.
RX   MEDLINE; 93190983.
RA   SHIPLEY J.M., KLINKENBERG M., WU B.M., BACHINSKY D.R., GRUBB J.H.,
RA   SLY W.S.;
RL   AM. J. HUM. GENET. 52:517-526(1993).
CC   -!- FUNCTION: BETA-GLUCURONIDASE PLAYS AN IMPORTANT ROLE IN THE
CC       DEGRADATION OF DERMATAN AND KERATAN SULFATES.
CC   -!- CATALYTIC ACTIVITY: A BETA-D-GLUCURONOSIDE + H(2)O = AN
CC       ALCOHOL + D-GLUCURONATE.
CC   -!- SUBUNIT: HOMOTETRAMER.
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- PTM: GLYCOSYLATED WITH 3 TO 4 N-LINKED OLIGOSACCHARIDE CHAINS.
CC   -!- DISEASE: DEFECTS IN GUSB ARE THE CAUSE OF MUCOPOLYSACCHARIDOSIS
CC       TYPE VII (MPS-VII) (ALSO KNOWN AS SLY SYNDROME).
CC   -!- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
DR   EMBL; M15182; G183233; -.
DR   EMBL; M65002; G183707; -.
DR   PIR; A26581; A26581.
DR   PDB; 1BHG; 17-SEP-97.
DR   MIM; 253220; -.
DR   PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR   PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
KW   HYDROLASE; GLYCOSIDASE; LYSOSOME; GLYCOPROTEIN; SIGNAL;
KW   MUCOPOLYSACCHARIDOSIS; DISEASE MUTATION; 3D-STRUCTURE.
FT   SIGNAL        1     22
FT   CHAIN        23    651       BETA-GLUCURONIDASE.
FT   ACT_SITE    451    451       PROTON DONOR.
FT   CARBOHYD    173    173
FT   CARBOHYD    272    272       POTENTIAL.
FT   CARBOHYD    420    420
FT   CARBOHYD    631    631       POTENTIAL.
FT   VARIANT     216    216       R -> W (IN MPS-VII).
FT   VARIANT     354    354       A -> V (IN MPS-VII).
FT   VARIANT     382    382       R -> C (IN MPS-VII).
FT   VARIANT     611    611       R -> W (IN MPS-VII).
FT   VARIANT     619    619       A -> V (IN MPS-VII).
FT   VARIANT     627    627       W -> C (IN MPS-VII).
SQ   SEQUENCE   651 AA;  74715 MW;  F0E4C8D6 CRC32;

  Query Match           61.3%;  Score 57;  DB 1;  Length 651;
  Best Local Similarity  63.6%;  Pred. No. 1.77e-01;
  Matches     7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    609 FTRQRQPKSAA 619
          |||||||  ::
Qy      3 FTRQRQPSEGS 13


RESULT    5
ID   BGLR_MOUSE     STANDARD;      PRT;   648 AA.
AC   P12265;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   BETA-GLUCURONIDASE PRECURSOR (EC 3.2.1.31).
GN   GUSB OR GUS.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88085188.
RA   GALLAGHER P.M., D'AMORE M.A., LUND S.D., ELLIOTT R.W., PAZIK J.,
RA   HOHMAN C., KORFHAGEN T.R., GANSCHOW R.E.;
RL   GENOMICS 1:145-152(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88284700.
RA   GALLAGHER P.M., D'AMORE M.A., LUND S.D., GANSCHOW R.E.;
RL   GENOMICS 2:215-219(1988).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89062453.
RA   D'AMORE M.A., GALLAGHER P.M., KORFHAGEN T.R., GANSCHOW R.E.;
RL   BIOCHEMISTRY 27:7131-7140(1988).
```

```
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89384641.
RA   WAWRZYNIAK C.J., GALLAGHER P.M., D'AMORE M.A., CARTER J.E.,
RA   LUND S.D., RINCHIK E.M., GANSCHOW R.E.;
RL   MOL. CELL. BIOL. 9:4074-4078(1989).
CC   -!- FUNCTION: BETA-GLUCURONIDASE PLAYS AN IMPORTANT ROLE IN THE
CC       DEGRADATION OF DERMATAN AND KERATAN SULFATES.
CC   -!- CATALYTIC ACTIVITY: A BETA-D-GLUCORONOSIDE + H(2)O = AN
CC       ALCOHOL + D-GLUCURONATE.
CC   -!- SUBUNIT: HOMOTETRAMER.
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
DR   EMBL; J03047; G309256; -.
DR   EMBL; J02836; G387180; -.
DR   EMBL; M63836; G193723; -.
DR   PIR; A28954; A28954.
DR   PIR; A29977; A29977.
DR   MGD; MGI:95874; GUS-S.
DR   PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR   PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
KW   HYDROLASE; GLYCOSIDASE; LYSOSOME; GLYCOPROTEIN; SIGNAL.
FT   SIGNAL       1     22
FT   CHAIN       23    648       BETA-GLUCURONIDASE.
FT   ACT_SITE   447    447       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD   172    172       POTENTIAL.
FT   CARBOHYD   416    416       POTENTIAL.
FT   CARBOHYD   591    591       POTENTIAL.
FT   CARBOHYD   627    627       POTENTIAL.
FT   CONFLICT   320    320       V -> I (IN REF. 4).
SQ   SEQUENCE   648 AA;  74239 MW;  13A10D2B CRC32;

  Query Match           60.2%;  Score 56;  DB 1;  Length 648;
  Best Local Similarity 100.0%;  Pred. No. 3.00e-01;
  Matches     7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    605 FTRQRQP 611
          |||||||
Qy      3 FTRQRQP 9


RESULT    6
ID   VPR_HV1Z2      STANDARD;      PRT;    96 AA.
AC   P12519;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN (R ORF PROTEIN).
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (Z2/CDC-Z34 ISOLATE) (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   THEODORE T., BUCKLER-WHITE A.;
RL   SUBMITTED (NOV-1988) TO THE HIV DATA BANK.
DR   EMBL; M22639; G329383; -.
DR   HIV; M22639; VPR$Z2Z6.
KW   AIDS.
SQ   SEQUENCE   96 AA;  11380 MW;  B28C76BE CRC32;

  Query Match           59.1%;  Score 55;  DB 1;  Length 96;
  Best Local Similarity 57.1%;  Pred. No. 5.07e-01;
  Matches     8;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGSS 94
          | :||||:: :|||
Qy      1 INFTRQRQPSEGSS 14


RESULT    7
ID   VPR_HV1MA      STANDARD;      PRT;    96 AA.
AC   P05955;
DT   01-NOV-1988 (REL. 09, CREATED)
DT   01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
DT   01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN (R ORF PROTEIN).
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (MAL ISOLATE) (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86245056.
RA   ALIZON M., WAIN-HOBSON S., MONTAGNIER L., SONIGO P.;
RL   CELL 46:63-74(1986).
DR   EMBL; K03456; G328024; -.
DR   EMBL; X04415; G60232; -.
DR   EMBL; A07116; G492875; -.
DR   HIV; K03456; VPR$MAL.
KW   AIDS.
SQ   SEQUENCE   96 AA;  11343 MW;  4AA5A84E CRC32;

  Query Match           59.1%;  Score 55;  DB 1;  Length 96;
  Best Local Similarity 57.1%;  Pred. No. 5.07e-01;
  Matches     8;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGSS 94
          | :||||:: :|||
Qy      1 INFTRQRQPSEGSS 14


RESULT    8
ID   SNF2_YEAST     STANDARD;      PRT;  1703 AA.
AC   P22082;
DT   01-AUG-1991 (REL. 19, CREATED)
DT   01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   TRANSCRIPTION REGULATORY PROTEIN SNF2 (SWI/SNF COMPLEX COMPONENT SNF2)
DE   (REGULATORY PROTEIN SWI2) (REGULATORY PROTEIN GAM1) (TRANSCRIPTION
DE   FACTOR TYE3).
GN   SNF2 OR SWI2 OR GAM1 OR TYE3 OR RIC1 OR YOR290C.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE; 91187857.
RA   LAURENT B.C., TREITEL M.A., CARLSON M.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 88:2687-2691(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AH22;
RX   MEDLINE; 91360076.
RA   YOSHIMOTO H., YAMASHITA I.;
RL   MOL. GEN. GENET. 228:270-280(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X2180-1B;
RX   MEDLINE; 95332261.
RA   KODAKI T., HOSAKA K., NIKAWA J., YAMASHITA S.;
RL   J. BIOCHEM. 117:362-368(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE; 97051594.
RA   CHERET G., BERNARDI A., SOR F.J.;
RL   YEAST 12:1059-1064(1996).
RN   [5]
RP   SEQUENCE OF 1-309 FROM N.A.
RC   STRAIN=S288C / FY1679;
RX   MEDLINE; 97298310.
RA   POIREY R., CZIEPLUCH C., TOBIASCH E., PUJOL A., KORDES E.,
RA   JAUNIAUX J.-C.;
RL   YEAST 13:479-482(1997).
RN   [6]
```

```
RP   MUTAGENESIS.
RX   MEDLINE; 97025355.
RA   RICHMOND E., PETERSON C.L.;
RL   NUCLEIC ACIDS RES. 24:3685-3692(1996).
CC   -!- FUNCTION: INVOLVED IN TRANSCRIPTIONAL ACTIVATION. THE SWI/SNF
CC       COMPLEX IS REQUIRED FOR THE INDUCED EXPRESSION OF A LARGE NUMBER
CC       OF GENES. THIS COMPLEX ALTERS CHROMATIN STRUCTURE TO FACILITATE
CC       BINDING OF GENE-SPECIFIC DEDICATED TRANSCRIPTION FACTORS.
CC   -!- SUBUNIT: COMPONENT OF THE SWI/SNF GLOBAL TRANSCRIPTION ACTIVATOR
CC       COMPLEX.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: STRONG, TO DROSOPHILA BRAHMA.
CC   -!- SIMILARITY: CONTAINS A COPY OF THE BROMODOMAIN.
CC   -!- SIMILARITY: TO HELICASES OF THE SNF2/RAD54 FAMILY.
DR   EMBL; M61703; G172632; -.
DR   EMBL; X57837; G4500; -.
DR   EMBL; D90459; G806532; -.
DR   EMBL; X89633; E189413; -.
DR   EMBL; Z75198; E252424; -.
DR   PIR; S15047; S15047.
DR   PIR; S16820; S16820.
DR   SGD; L0001945; SNF2.
DR   PROSITE; PS00633; BROMODOMAIN_1; 1.
DR   PROSITE; PS50014; BROMODOMAIN_2; 1.
KW   TRANSCRIPTION REGULATION; NUCLEAR PROTEIN; ACTIVATOR; REPEAT;
KW   BROMODOMAIN; ATP-BINDING; HELICASE.
FT   DOMAIN       55     68       GLN-RICH.
FT   DOMAIN      207    239       ALA/GLN-RICH.
FT   NP_BIND     792    799       ATP (BY SIMILARITY).
FT   SITE        894    897       DEGH BOX.
FT   DOMAIN     1446   1458       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN     1505   1522       9 X 2 AA TANDEM REPEATS OF R-G.
FT   DOMAIN     1568   1638       BROMODOMAIN.
SQ   SEQUENCE   1703 AA;  194050 MW;  CA278760 CRC32;

  Query Match           59.1%;  Score 55;  DB 1;  Length 1703;
  Best Local Similarity  50.0%;  Pred. No. 5.07e-01;
  Matches     5;  Conservative     5;  Mismatches   0;  Indels   0;  Gaps   0;

Db    299 EFARRRQPTD 308
          :|:|:|||::
Qy      2 NFTRQRQPSE 11


RESULT    9
ID   YN44_YEAST     STANDARD;      PRT;   355 AA.
AC   P48559;
DT   01-FEB-1996 (REL. 33, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 40.7 KD GTP-BINDING PROTEIN IN MCK1-RP55B INTERGENIC
DE   REGION.
GN   YNL304W OR N0410.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RX   MEDLINE; 96132033.
RA   MAURER K.C.T., URBANUS J.H.M., PLANTA R.J.;
RL   YEAST 11:1303-1310(1995).
DR   EMBL; U23084; G1050856; -.
DR   EMBL; Z71580; E239733; -.
KW   HYPOTHETICAL PROTEIN; GTP-BINDING; LIPOPROTEIN; PRENYLATION.
FT   NP_BIND      35     42       GTP (POTENTIAL).
FT   NP_BIND     166    170       GTP (POTENTIAL).
FT   NP_BIND     230    233       GTP (POTENTIAL).
FT   LIPID       353    353       GERANYL-GERANYL (BY SIMILARITY).
FT   LIPID       354    354       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   355 AA;  40678 MW;  9E3B3359 CRC32;

  Query Match           58.1%;  Score 54;  DB 1;  Length 355;
  Best Local Similarity  87.5%;  Pred. No. 8.49e-01;
  Matches     7;  Conservative     1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    321 FNFTRQRQ 328
          :|||||||
Qy      1 INFTRQRQ 8


RESULT   10
ID   VPR_HV1OY      STANDARD;      PRT;    96 AA.
AC   P20891;
DT   01-FEB-1991 (REL. 17, CREATED)
DT   01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT   01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN (R ORF PROTEIN).
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (OYI ISOLATE) (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90148544.
RA   HUET T., DAZZA M.C., BRUN-VEZINET F., ROELANTS G.E., WAIN-HOBSON S.;
RL   AIDS 3:707-715(1989).
CC   -!- THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A HEALTHY GABONESE
CC       INDIVIDUAL.
DR   EMBL; M26727; G328446; -.
DR   HIV; M26727; VPR$OYI.
KW   AIDS.
SQ   SEQUENCE   96 AA;  11494 MW;  4C01E21D CRC32;

  Query Match           57.0%;  Score 53;  DB 1;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 1.41e+00;
  Matches     7;  Conservative     5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   11
ID   VPR_HV1JR      STANDARD;      PRT;    96 AA.
AC   P20883;
DT   01-FEB-1991 (REL. 17, CREATED)
DT   01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT   01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN (R ORF PROTEIN).
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (JRCSF ISOLATE) (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   KOYANAGI S., CHEN I.S.Y.;
RL   SUBMITTED (DEC-1988) TO THE HIV DATA BANK.
DR   EMBL; M38429; G327817; -.
DR   HIV; M38429; VPR$JRCSF.
KW   AIDS.
SQ   SEQUENCE   96 AA;  11419 MW;  1DC76121 CRC32;

  Query Match           57.0%;  Score 53;  DB 1;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 1.41e+00;
  Matches     7;  Conservative     5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   12
ID   VPR_HV1RH      STANDARD;      PRT;    96 AA.
AC   P05954;
DT   01-NOV-1988 (REL. 09, CREATED)
DT   01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
```

```
DT   01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN (R ORF PROTEIN).
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (RF/HAT ISOLATE) (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   STARCICH B.R., HAHN B.H., SHAW G.M., MCNEELY P.D., MODROW S.,
RA   WOLF H., PARKS E.S., PARKS W.P., JOSEPHS S.F., GALLO R.C.,
RA   WONG-STAAL F.;
RL   SUBMITTED (XXX-1987) TO THE HIV DATA BANK.
DR   EMBL; M17451; G328571; -.
DR   HIV; M17451; VPR$RF.
KW   AIDS.
SQ   SEQUENCE   96 AA;  11338 MW;  B2F633A3 CRC32;

  Query Match           57.0%;  Score 53;  DB 1;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 1.41e+00;
  Matches    7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   13
ID   ACH5_CAEEL     STANDARD;      PRT;   511 AA.
AC   Q23022; Q17408; P91265; O02559;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   ACETYLCHOLINE RECEPTOR LIKE PROTEIN, ALPHA-TYPE SUBUNIT UNC-38
DE   PRECURSOR.
GN   UNC-38 OR F21F3.5.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   FLEMING J.T., SQUIRE M.D., BARNES T.M., TORNOE C., MATSUDA K.,
RA   SULSTON J.E., BARNARD E.A., SATTELLE D.B., LEWIS J.A.;
RL   SUBMITTED (DEC-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   GEISEL C., KRAMER J., ELLIOTT G.;
RL   SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: POSSIBLE ACETYLCHOLINE RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR   EMBL; X98600; E285245; -.
DR   EMBL; X98599; E285244; -.
DR   EMBL; U88175; G1825663; -.
DR   WORMPEP; F21F3.5; CE09535.
DR   PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW   RECEPTOR; POSTSYNAPTIC MEMBRANE; IONIC CHANNEL; GLYCOPROTEIN;
KW   TRANSMEMBRANE; SIGNAL.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17    511       ACETYLCHOLINE RECEPTOR LIKE PROTEIN,
FT                                ALPHA-TYPE SUBUNIT UNC-38.
FT   DOMAIN       17    261       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    262    282       POTENTIAL.
FT   TRANSMEM    291    311       POTENTIAL.
FT   TRANSMEM    324    344       POTENTIAL.
FT   DOMAIN      345    464       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    465    485       POTENTIAL.
FT   DISULFID    151    165       BY SIMILARITY.
FT   DISULFID    238    239       BY SIMILARITY.
FT   CARBOHYD    124    124       POTENTIAL.
FT   CARBOHYD    202    202       POTENTIAL.
FT   CONFLICT    248    248       T -> TKHEKLQSLNKKLQPNFARIEKYI (IN
FT                                REF. 2).
SQ   SEQUENCE   511 AA;  59454 MW;  E3009E18 CRC32;

  Query Match           57.0%;  Score 53;  DB 1;  Length 511;
  Best Local Similarity  35.7%;  Pred. No. 1.41e+00;
  Matches    5;  Conservative    7;  Mismatches   2;  Indels   0;  Gaps   0;

Db     36 VNYNRHRRPSTSPN 49
          :|: |:|:|| :::
Qy      1 INFTRQRQPSEGSS 14


RESULT   14
ID   TOXA_PSEAE     STANDARD;      PRT;   638 AA.
AC   P11439;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   EXOTOXIN A PRECURSOR (NAD-DEPENDENT ADP-RIBOSYLTRANSFERASE
DE   (EC 2.4.2.-)).
GN   ETA.
OS   PSEUDOMONAS AERUGINOSA.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   PSEUDOMONADACEAE.
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 26-53.
RX   MEDLINE; 84194063.
RA   GRAY G.L., SMITH D.H., BALDRIDGE J.S., HARKINS R.N., VASIL M.L.,
RA   CHEN E.Y., HEYNEKER H.L.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 81:2645-2649(1984).
RN   [2]
RP   ACTIVE SITE.
RX   MEDLINE; 87250491.
RA   CARROLL S.F., COLLIER R.J.;
RL   J. BIOL. CHEM. 262:8707-8711(1987).
RN   [3]
RP   DOMAINS.
RX   MEDLINE; 90375493.
RA   CHAUDHARY V.K., JINNO Y., GALO M.G., FITZGERALD D., PASTAN I.;
RL   J. BIOL. CHEM. 265:16306-16310(1990).
RN   [4]
RP   DOMAINS.
RX   MEDLINE; 91006124.
RA   BOURDENET S., VACHERON M.-J., GUINAND M., MICHEL G., ARMINJON F.;
RL   EUR. J. BIOCHEM. 192:379-385(1990).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 424-638.
RX   MEDLINE; 96016159.
RA   LI M., DYDA F., BENHAR I., PASTAN I., DAVIES D.R.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 92:9308-9312(1995).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 424-638.
RX   MEDLINE; 96293446.
RA   LI M., DYDA F., BENHAR I., PASTAN I., DAVIES D.R.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 93:6902-6906(1996).
CC   -!- FUNCTION: THIS TOXIN IS AN NAD-DEPENDENT ADP-RIBOSYLTRANSFERASE.
CC       IT CATALYZES THE TRANSFER OF THE ADP RIBOSYL MOIETY OF OXIDIZED
CC       NAD ONTO ELONGATION FACTOR 2 (EF-2) THUS ARRESTING PROTEIN
CC       SYNTHESIS.
CC   -!- PTM: THE 8 CYSTEINES PARTICIPATE IN INTRACHAIN DISULFIDE BONDS.
CC   -!- SIMILARITY: REGIONAL SEQUENCE SIMILARITY AT THE ACTIVE SITE
CC       WITH DIPHTHERIA TOXIN (DT).
DR   EMBL; K01397; G151216; -.
DR   PIR; A30347; A30347.
DR   PDB; 1AER; 10-JUN-96.
DR   PDB; 1DMA; 15-SEP-95.
KW   TOXIN; SIGNAL; TRANSFERASE; GLYCOSYLTRANSFERASE; NAD; 3D-STRUCTURE.
FT   SIGNAL        1     25
FT   CHAIN        26    638       EXOTOXIN A.
FT   DOMAIN       26    277       IA (REQUIRED FOR TARGET CELL
FT                                RECOGNITION).
FT   DOMAIN      278    389       II (REQUIRED FOR TRANSLOCATION IN TARGET
FT                                CELL CYTOPLASM).
FT   DOMAIN      390    429       IB.
```

```
FT   DOMAIN      430    638       III (REQUIRED FOR ADP-RIBSOSYL ACTIVITY).
FT   ACT_SITE    465    465
FT   ACT_SITE    578    578       INTERACT WITH NAD.
SQ   SEQUENCE   638 AA;  69309 MW;  33D9876A CRC32;

  Query Match           57.0%;  Score 53;  DB 1;  Length 638;
  Best Local Similarity  85.7%;  Pred. No. 1.41e+00;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    297 FTRHRQP 303
          |||:|||
Qy      3 FTRQRQP 9


RESULT   15
ID   SRA9_CAEEL     STANDARD;      PRT;   331 AA.
AC   Q09212;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   SRA-9 PROTEIN.
GN   SRA-9 OR AH6.14.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   JASSAL B.;
RL   SUBMITTED (JAN-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE SRA FAMILY OF C.ELEGANS RECEPTOR-LIKE
CC       PROTEINS.
DR   EMBL; Z48009; G643104; -.
DR   WORMPEP; AH6.14; CE01455.
KW   TRANSMEMBRANE; MULTIGENE FAMILY.
FT   TRANSMEM     26     46       POTENTIAL.
FT   TRANSMEM    104    124       POTENTIAL.
FT   TRANSMEM    143    163       POTENTIAL.
FT   TRANSMEM    189    209       POTENTIAL.
FT   TRANSMEM    238    258       POTENTIAL.
FT   TRANSMEM    275    295       POTENTIAL.
SQ   SEQUENCE   331 AA;  38284 MW;  E9D38381 CRC32;

  Query Match           54.8%;  Score 51;  DB 1;  Length 331;
  Best Local Similarity  54.5%;  Pred. No. 3.83e+00;
  Matches     6;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    310 IEFTKQKQSQE 320
          |:||:|:|: |
Qy      1 INFTRQRQPSE 11


RESULT   16
ID   VPR_HV1EL      STANDARD;      PRT;    96 AA.
AC   P05956;
DT   01-NOV-1988 (REL. 09, CREATED)
DT   01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
DT   01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN (R ORF PROTEIN).
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (ELI ISOLATE) (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86245056.
RA   ALIZON M., WAIN-HOBSON S., MONTAGNIER L., SONIGO P.;
RL   CELL 46:63-74(1986).
DR   EMBL; K03454; G326681; -.
DR   EMBL; A07108; G492867; -.
DR   HIV; K03454; VPR$ELI.
KW   AIDS.
SQ   SEQUENCE    96 AA;  11306 MW;  FE489F84 CRC32;

  Query Match           53.8%;  Score 50;  DB 1;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 6.23e+00;
  Matches     7;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db     81 IGIIRQRRARNGSS 94
          | : |||:: :|||
Qy      1 INFTRQRQPSEGSS 14


RESULT   17
ID   FBP_NEIGO      STANDARD;      PRT;   330 AA.
AC   P17259;
DT   01-AUG-1990 (REL. 15, CREATED)
DT   01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT   01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE   MAJOR FERRIC IRON BINDING PROTEIN PRECURSOR (FBP) (MAJOR IRON-
DE   REGULATED PROTEIN) (MIRP).
GN   FBP.
OS   NEISSERIA GONORRHOEAE.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   NEISSERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=F62;
RX   MEDLINE; 90237747.
RA   BERISH S.A., MIETZNER T.A., MAYER L.W., GENCO C.A., HOLLOWAY B.P.,
RA   MORSE S.A.;
RL   J. EXP. MED. 171:1535-1546(1990).
RN   [2]
RP   SEQUENCE OF 23-69.
RX   MEDLINE; 87168188.
RA   MIETZNER T.A., BOLAN G., SCHOOLNIK G.K., MORSE S.A.;
RL   J. EXP. MED. 165:1041-1057(1987).
CC   -!- FUNCTION: THIS PROTEIN MAY BE A CENTRAL COMPONENT IN THE IRON-
CC       ACQUISITION SYSTEM.
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC   -!- INDUCTION: UNDER IRON-DEFICIENCY.
CC   -!- IRON CO-PURIFIES WITH FBP AND IS BOUND BY THE PROTEIN AS A
CC       FERRIC ION AT AN APPROXIMATE MOLAR RATIO OF 1:1.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC       PROTEIN FAMILY 1.
DR   EMBL; X51901; G44858; -.
DR   PIR; S10256; S10256.
DR   PROSITE; PS01037; SBP_BACTERIAL_1; 1.
KW   TRANSPORT; IRON TRANSPORT; PERIPLASMIC; SIGNAL.
FT   SIGNAL        1     22
FT   CHAIN        23    330       FERRIC IRON BINDING PROTEIN.
SQ   SEQUENCE   330 AA;  35769 MW;  F18F3562 CRC32;

  Query Match           53.8%;  Score 50;  DB 1;  Length 330;
  Best Local Similarity  50.0%;  Pred. No. 6.23e+00;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    235 LNFVRHRDPG 244
          :|| |:|:|:
Qy      1 INFTRQRQPS 10


RESULT   18
ID   FBP_NEIME      STANDARD;      PRT;   330 AA.
AC   P17940;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE   MAJOR FERRIC IRON BINDING PROTEIN PRECURSOR (FBP) (MAJOR IRON-
DE   REGULATED PROTEIN) (MIRP).
GN   FBP.
OS   NEISSERIA MENINGITIDIS.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   NEISSERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=80084313;
RX   MEDLINE; 90356404.
RA   BERISH S.A., KAPCZYNSKI D.R., MORSE S.A.;
RL   NUCLEIC ACIDS RES. 18:4596-4596(1990).
CC   -!- FUNCTION: THIS PROTEIN MAY BE A CENTRAL COMPONENT IN THE IRON-
CC       ACQUISITION SYSTEM.
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC   -!- INDUCTION: UNDER IRON-DEFICIENCY.
CC   -!- IRON CO-PURIFIES WITH FBP AND IS BOUND BY THE PROTEIN AS A
CC       FERRIC ION AT AN APPROXIMATE MOLAR RATIO OF 1:1.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC       PROTEIN FAMILY 1.
DR   EMBL; X53467; G45038; -.
DR   PIR; S10978; S10978.
DR   PIR; B60816; B60816.
DR   PROSITE; PS01037; SBP_BACTERIAL_1; 1.
KW   TRANSPORT; IRON TRANSPORT; PERIPLASMIC; SIGNAL.
FT   SIGNAL        1     22
FT   CHAIN        23    330       FERRIC IRON BINDING PROTEIN.
SQ   SEQUENCE   330 AA;  35741 MW;  AD87830D CRC32;

  Query Match           53.8%;  Score 50;  DB 1;  Length 330;
  Best Local Similarity  50.0%;  Pred. No. 6.23e+00;
  Matches    5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    235 LNFVRHRDPG 244
          :|| |:|:|:
Qy      1 INFTRQRQPS 10


RESULT   19
ID   CLPA_RHOBL     STANDARD;      PRT;   793 AA.
AC   P05444;
DT   01-NOV-1988 (REL. 09, CREATED)
DT   01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   CLPA HOMOLOG PROTEIN.
OS   RHODOPSEUDOMONAS BLASTICA.
OC   PROKARYOTA; GRACILICUTES; ANOXYPHOTOBACTERIA; PURPLE BACTERIA;
OC   RHODOSPIRILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 85058188.
RA   TYBULEWICZ V.L.J., FALK G., WALKER J.E.;
RL   J. MOL. BIOL. 179:185-214(1984).
CC   -!- SIMILARITY: HIGH, TO OTHER MEMBERS OF THE CLPA/CLPB FAMILY.
DR   EMBL; Z00018; G45972; -.
DR   PIR; S04667; SURFCA.
DR   HSSP; P02392; 1CTF.
DR   PROSITE; PS00870; CLPAB_1; 1.
DR   PROSITE; PS00871; CLPAB_2; 1.
KW   CHAPERONE; ATP-BINDING; REPEAT.
FT   DOMAIN      199    447       I.
FT   DOMAIN      451    639       II.
FT   NP_BIND     244    251       ATP (POTENTIAL).
FT   NP_BIND     525    532       ATP (POTENTIAL).
SQ   SEQUENCE   793 AA;  86891 MW;  F50F0D18 CRC32;

  Query Match           53.8%;  Score 50;  DB 1;  Length 793;
  Best Local Similarity  35.7%;  Pred. No. 6.23e+00;
  Matches    5;  Conservative    5;  Mismatches   4;  Indels   0;  Gaps   0;

Db    647 IGFGRERRTGEDTA 660
          | | |:|: :| ::
Qy      1 INFTRQRQPSEGSS 14


RESULT   20
ID   TRK2_YEAST     STANDARD;      PRT;   889 AA.
AC   P28584;
DT   01-DEC-1992 (REL. 24, CREATED)
DT   01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   POTASSIUM TRANSPORT PROTEIN, LOW-AFFINITY.
GN   TRK2 OR RPD2 OR YKR050W.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91304419.
RA   KO C.H., GABER R.F.;
RL   MOL. CELL. BIOL. 11:4266-4273(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   VISSERS S., URRESTARAZU L.A., JAUNIAUX J.-C.;
RL   SUBMITTED (MAR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THIS PROTEIN IS REQUIRED FOR LOW-AFFINITY POTASSIUM
CC       TRANSPORT.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
DR   EMBL; M65215; G173033; -.
DR   EMBL; Z28275; G486503; -.
DR   PIR; A41259; A41259.
DR   SGD; L0002335; TRK2.
KW   TRANSPORT; TRANSMEMBRANE; POTASSIUM TRANSPORT; GLYCOPROTEIN.
FT   TRANSMEM     52     73       POTENTIAL.
FT   TRANSMEM     81    101       POTENTIAL.
FT   TRANSMEM    110    130       POTENTIAL.
FT   TRANSMEM    465    487       POTENTIAL.
FT   TRANSMEM    500    521       POTENTIAL.
FT   TRANSMEM    525    545       POTENTIAL.
FT   TRANSMEM    549    569       POTENTIAL.
FT   TRANSMEM    585    605       POTENTIAL.
FT   TRANSMEM    610    630       POTENTIAL.
FT   TRANSMEM    658    678       POTENTIAL.
FT   TRANSMEM    744    764       POTENTIAL.
FT   TRANSMEM    777    797       POTENTIAL.
FT   CARBOHYD    216    216       POTENTIAL.
FT   CARBOHYD    233    233       POTENTIAL.
FT   CARBOHYD    265    265       POTENTIAL.
FT   CARBOHYD    606    606       POTENTIAL.
FT   CARBOHYD    700    700       POTENTIAL.
FT   CARBOHYD    801    801       POTENTIAL.
SQ   SEQUENCE   889 AA;  101086 MW;  F960B401 CRC32;

  Query Match           53.8%;  Score 50;  DB 1;  Length 889;
  Best Local Similarity  60.0%;  Pred. No. 6.23e+00;
  Matches    6;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    233 NFSSKRQPSD 242
          ||:  ||||:
Qy      2 NFTRQRQPSE 11


RESULT   21
ID   VPR_HV1N5      STANDARD;      PRT;    96 AA.
AC   P12520;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN (R ORF PROTEIN).
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (NEW YORK-5 ISOLATE) (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   BUCKLER C.E., BUCKLER-WHITE A.J., WILLEY R.L., MCCOY J.;
RL   SUBMITTED (JUN-1988) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; M19921; G328421; -.
DR   HIV; M19921; VPR$NL43.
KW   AIDS.
SQ   SEQUENCE   96 AA;  11379 MW;  FF4FB7C0 CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 1.01e+01;
  Matches    7;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;
```

```
Db      81 IGVTRQRRARNGAS 94
           |  ||||:: :|:|
Qy       1 INFTRQRQPSEGSS 14


RESULT   22
ID   RNP_HYDHY      STANDARD;      PRT;   128 AA.
AC   P00677;
DT   21-JUL-1986 (REL. 01, CREATED)
DT   21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   RIBONUCLEASE PANCREATIC (EC 3.1.27.5) (RNASE 1) (RNASE A).
GN   RNASE1 OR RNS1.
OS   HYDROCHOERUS HYDROCHAERIS (CAPYBARA) (CARPINCHO).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 87036770.
RA   BEINTEMA J.J., NEUTEBOOM B.;
RL   J. MOL. EVOL. 19:145-152(1983).
CC   -!- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 3'-PHOSPHOMONO-
CC       NUCLEOTIDES AND 3'-PHOSPHOOLIGONUCLEOTIDES ENDING IN C-P OR U-P
CC       WITH 2',3'-CYCLIC PHOSPHATE INTERMEDIATES.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- TISSUE SPECIFICITY: PANCREAS.
CC   -!- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
DR   PIR; A00824; NRYY.
DR   HSSP; P00656; 1SRN.
DR   PROSITE; PS00127; RNASE_PANCREATIC; 1.
KW   HYDROLASE; NUCLEASE; ENDONUCLEASE.
FT   DISULFID     26     84       BY SIMILARITY.
FT   DISULFID     40     95       BY SIMILARITY.
FT   DISULFID     58    110       BY SIMILARITY.
FT   DISULFID     65     72       BY SIMILARITY.
FT   ACT_SITE     12     12       BY SIMILARITY.
FT   ACT_SITE     41     41       BY SIMILARITY.
FT   ACT_SITE    119    119       BY SIMILARITY.
SQ   SEQUENCE   128 AA;  14345 MW;  23DFF27D CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 128;
  Best Local Similarity  57.1%;  Pred. No. 1.01e+01;
  Matches     8;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db       6 MKFQRQHVDSEGSS 19
           ::| ||:  |||||
Qy       1 INFTRQRQPSEGSS 14


RESULT   23
ID   RS7_MYCSM      STANDARD;      PRT;   155 AA.
AC   P41193;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   30S RIBOSOMAL PROTEIN S7.
GN   RPSG.
OS   MYCOBACTERIUM SMEGMATIS.
OC   PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LR222;
RX   MEDLINE; 95014056.
RA   KENNEY T.J., CHURCHWARD G.;
RL   J. BACTERIOL. 176:6153-6156(1994).
CC   -!- FUNCTION: PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF
CC       16S RIBOSOMAL RNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.
DR   EMBL; L34681; G511652; -.
DR   PROSITE; PS00052; RIBOSOMAL_S7; 1.
KW   RIBOSOMAL PROTEIN; RRNA-BINDING.
FT   INIT_MET      0      0       BY SIMILARITY.
SQ   SEQUENCE   155 AA;  17497 MW;  0409A286 CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 155;
  Best Local Similarity  62.5%;  Pred. No. 1.01e+01;
  Matches     5;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db     104 VNFSRQRR 111
           :||:|||:
Qy       1 INFTRQRQ 8


RESULT   24
ID   NRL_MOUSE      STANDARD;      PRT;   237 AA.
AC   P54846;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   NEURAL RETINA-SPECIFIC LEUCINE ZIPPER PROTEIN (NRL).
GN   NRL.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=RETINA;
RX   MEDLINE; 94116986.
RA   FARJO Q., JACKSON A.U., XU J., GRYZENIA M., SKOLNICK C.,
RA   AGARWAL N., SWAROOP A.;
RL   GENOMICS 18:216-222(1993).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: NEURAL RETINA.
CC   -!- SIMILARITY: TO OTHER BZIP PROTEINS.
DR   EMBL; L14935; G388917; -.
DR   TRANSFAC; T01438; -.
DR   MGD; MGI:102567; NRL.
KW   TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN.
FT   DNA_BIND    159    185       BASIC MOTIF.
FT   DOMAIN      187    208       LEUCINE-ZIPPER.
SQ   SEQUENCE   237 AA;  26083 MW;  8125334E CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 237;
  Best Local Similarity  50.0%;  Pred. No. 1.01e+01;
  Matches     7;  Conservative    3;  Mismatches   4;  Indels   0;  Gaps   0;

Db      19 MKFEIKREPSEGRS 32
           ::|   |:|||| |
Qy       1 INFTRQRQPSEGSS 14


RESULT   25
ID   MAG2_HUMAN     STANDARD;      PRT;   314 AA.
AC   P43356;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 2 (MAGE-2 ANTIGEN).
GN   MAGEA2 OR MAGE2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94102805.
RA   DE SMET C., LURQUIN C., VAN DER BRUGGEN P., DE PLAEN E., BRASSEUR F.,
RA   BOON T.;
RL   IMMUNOGENETICS 39:121-129(1994).
RN   [2]
RP   MUTAGENESIS.
RC   TISSUE=BLOOD;
RX   MEDLINE; 94157413.
RA   GAUGLER B., VAN DEN EYNDE B., VAN DER BRUGGEN P., ROMERO P.,
RA   GAFORIO J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;
RL   J. EXP. MED. 179:921-930(1994).
```

```
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC       CYTOLYTIC T LYMPHOCYTES.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY (90%)
CC       WITH MAGE-12.
DR   EMBL; L18920; G436181; -.
KW   ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT   DOMAIN       40     43       POLY-SER.
FT   MUTAGEN     170    170       V->D: IMPROVES ABILITY TO BIND TO HLA-A1.
SQ   SEQUENCE   314 AA;  35055 MW;  E101F41E CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 314;
  Best Local Similarity  64.3%;  Pred. No. 1.01e+01;
  Matches    9;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db     75 INYTLWRQSDEGSS 88
          ||:|  ||: ||||
Qy      1 INFTRQRQPSEGSS 14


RESULT   26
ID   NRLB_KLEPN     STANDARD;      PRT;   349 AA.
AC   P10045;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT   01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE   NITRILASE, BROMOXYNIL-SPECIFIC (EC 3.5.5.1).
GN   BXN.
OS   KLEBSIELLA PNEUMONIAE.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88198177.
RA   STALKER D.M., MALYJ L.D., MCBRIDE K.E.;
RL   J. BIOL. CHEM. 263:6310-6314(1988).
CC   -!- FUNCTION: SPECIFIC FOR THE HERBICIDE BROMOXYNIL.
CC   -!- CATALYTIC ACTIVITY: A NITRILE + H(2)O = A CARBOXYLATE + NH(3).
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- EXPRESSION OF THIS BACTERIAL NITRILASE IN TRANSGENIC PLANTS
CC       RESULTS IN HIGH LEVELS OF RESISTANCE TO THE HERBICIDE BROMOXYNIL.
CC   -!- SIMILARITY: TO OTHER NITRILASES.
DR   EMBL; J03196; G149175; -.
DR   PIR; A28658; A28658.
DR   PROSITE; PS00920; NITRIL_CHT_1; 1.
DR   PROSITE; PS00921; NITRIL_CHT_2; 1.
KW   HYDROLASE; HERBICIDE RESISTANCE.
FT   ACT_SITE    161    161       BY SIMILARITY.
SQ   SEQUENCE   349 AA;  37801 MW;  5FEE4AFA CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 349;
  Best Local Similarity  50.0%;  Pred. No. 1.01e+01;
  Matches    5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    298 VSINRQRQPA 307
          ::: |||||:
Qy      1 INFTRQRQPS 10


RESULT   27
ID   SCRB_STAXY     STANDARD;      PRT;   494 AA.
AC   Q05936;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   SUCROSE-6-PHOSPHATE HYDROLASE (EC 3.2.1.26) (SUCRASE) (INVERTASE).
GN   SCRB.
OS   STAPHYLOCOCCUS XYLOSUS.
OC   PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93139055.
RA   BRUECKNER R., WAGNER E., GOETZ F.;
RL   J. BACTERIOL. 175:851-857(1993).
CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA-D-
CC       FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.
CC   -!- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
DR   EMBL; X67744; E264653; -.
DR   PIR; A47059; A47059.
DR   PROSITE; PS00609; GLYCOSYL_HYDROL_F32; 1.
KW   HYDROLASE; GLYCOSIDASE.
FT   ACT_SITE     48     48       BY SIMILARITY.
SQ   SEQUENCE   494 AA;  57371 MW;  8F97236A CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 494;
  Best Local Similarity  63.6%;  Pred. No. 1.01e+01;
  Matches    7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    358 KFTRQLHPYEG 368
          :|||| :| ||
Qy      2 NFTRQRQPSEG 12


RESULT   28
ID   MET7_NEUCR     STANDARD;      PRT;   542 AA.
AC   P38675;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   CYSTATHIONINE GAMMA-SYNTHASE (EC 4.2.99.9) (O-SUCCINYLHOMOSERINE
DE   (THIOL)-LYASE).
GN   MET-7.
OS   NEUROSPORA CRASSA.
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; PYRENOMYCETES; SORDARIALES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92175534.
RA   CRAWFORD J.M., GEEVER R.F., ASCH D.K., CASE M.E.;
RL   GENE 111:265-266(1992).
CC   -!- CATALYTIC ACTIVITY: O-SUCCINYL-L-HOMOSERINE + L-CYSTEINE =
CC       CYSTATHIONINE + SUCCINATE (CAN ALSO USE HYDROGEN SULFIDE AND
CC       METHANETHIOL AS SUBSTRATES).
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC   -!- PATHWAY: SECOND STEP IN METHIONINE BIOSYNTHESIS.
CC   -!- SUBUNIT: MET-3 AND MET-7 ARE REQUIRED TO FORM A FUNCTIONAL
CC       CYSTATHIONINE GAMMA-SYNTHASE.
CC   -!- SIMILARITY: STRONG, TO YEAST YJR130C AND YML082W.
DR   EMBL; M64066; -; NOT_ANNOTATED_CDS.
DR   PIR; JQ1524; JQ1524.
KW   METHIONINE BIOSYNTHESIS; LYASE; PYRIDOXAL PHOSPHATE.
FT   BINDING     350    350       PYRIDOXAL PHOSPHATE (POTENTIAL).
SQ   SEQUENCE   542 AA;  60196 MW;  A23C9D9D CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 542;
  Best Local Similarity  36.4%;  Pred. No. 1.01e+01;
  Matches    4;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db     13 VDFIRSRAPAD 23
          ::| | | |::
Qy      1 INFTRQRQPSE 11


RESULT   29
ID   BGLR_RAT       STANDARD;      PRT;   648 AA.
AC   P06760;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE   BETA-GLUCURONIDASE PRECURSOR (EC 3.2.1.31).
GN   GUS.
OS   RATTUS NORVEGICUS (RAT).
```

```
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PREPUTIAL GLAND;
RX   MEDLINE; 87016933.
RA   NISHIMURA Y., ROSENFELD M.G., KREIBICH G., GUBLER U., SABATINI D.D.,
RA   ADESNIK M., ANDY R.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:7292-7296(1986).
RN   [2]
RP   SEQUENCE OF 14-648 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 88183378.
RA   POWELL P.P., KYLE J.W., MILLER R.D., PANTANO J., GRUBB J.H., SLY W.S.;
RL   BIOCHEM. J. 250:547-555(1988).
CC   -!- FUNCTION: BETA-GLUCURONIDASE PLAYS AN IMPORTANT ROLE IN THE
CC       DEGRADATION OF DERMATAN AND KERATAN SULFATES.
CC   -!- CATALYTIC ACTIVITY: A BETA-D-GLUCURONOSIDE + H(2)O = AN
CC       ALCOHOL + D-GLUCURONATE.
CC   -!- SUBUNIT: HOMOTETRAMER.
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- PTM: UNDERGOES A POSTTRANSCRIPTIONAL PROTEOLYTIC CLEAVAGE NEAR ITS
CC       C-TERMINAL END, WHICH REDUCES ITS SIZE BY APPROXIMATELY 3 KD. THE
CC       SITE OF THIS CLEAVAGE HAS AS YET NOT BEEN DETERMINED.
CC   -!- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
DR   EMBL; M13962; G204330; -.
DR   EMBL; Y00717; G758260; -.
DR   PIR; A25047; A25047.
DR   PIR; S00345; S00345.
DR   PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR   PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
KW   HYDROLASE; GLYCOSIDASE; LYSOSOME; GLYCOPROTEIN; SIGNAL.
FT   SIGNAL       1    22
FT   CHAIN       23   648       BETA-GLUCURONIDASE.
FT   ACT_SITE   447   447       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD   172   172       POTENTIAL.
FT   CARBOHYD   416   416       POTENTIAL.
FT   CARBOHYD   591   591       POTENTIAL.
FT   CARBOHYD   627   627       POTENTIAL.
FT   CONFLICT    14    14       Q -> E (IN REF. 2).
FT   CONFLICT    21    21       V -> L (IN REF. 2).
FT   CONFLICT   487   487       M -> L (IN REF. 2).
SQ   SEQUENCE   648 AA;  74793 MW;  AF91C615 CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 648;
  Best Local Similarity  85.7%;  Pred. No. 1.01e+01;
  Matches     6;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Db    605 FTRQRNP 611
          ||||| |
Qy      3 FTRQRQP 9


RESULT   30
ID   ERG7_RAT       STANDARD;      PRT;   733 AA.
AC   P48450;
DT   01-FEB-1996 (REL. 33, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   LANOSTEROL SYNTHASE (EC 5.4.99.7) (OXIDOSQUALENE--LANOSTEROL CYCLASE)
DE   (2,3-EPOXYSQUALENE--LANOSTEROL CYCLASE) (OSC).
GN   OSC OR LSS.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX   MEDLINE; 95253156.
RA   KUSANO M., SHIBUYA M., SANKAWA U., EBIZUKA Y.;
RL   BIOL. PHARM. BULL. 18:195-197(1995).
CC   -!- FUNCTION: CATALYZES THE CYCLIZATION OF (S)-2,3 OXIDOSQUALENE TO
CC       LANOSTEROL, A REACTION THAT FORMS THE STEROL NUCLEUS.
CC   -!- CATALYTIC ACTIVITY: (S)-2,3-EPOXYSQUALENE = LANOSTEROL.
CC   -!- PATHWAY: INITIAL STEP IN CHOLESTEROL, STEROID HORMONES AND VITAMIN
CC       D BIOSYNTHESIS.
CC   -!- SIMILARITY: BELONGS TO THE FAMILY OF TERPENE CYCLASES/MUTASES.
DR   EMBL; D45252; G639865; -.
DR   PROSITE; PS01074; TERPENE_SYNTHASES; 1.
KW   ISOMERASE; STEROID BIOSYNTHESIS.
SQ   SEQUENCE   733 AA;  83305 MW;  6ED7B647 CRC32;

  Query Match           52.7%;  Score 49;  DB 1;  Length 733;
  Best Local Similarity  38.5%;  Pred. No. 1.01e+01;
  Matches     5;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db    565 LDFCRKKQRADGS 577
          ::| | :| ::||
Qy      1 INFTRQRQPSEGS 13


RESULT   31
ID   RT19_PETHY     STANDARD;      PRT;    94 AA.
AC   P27527;
DT   01-AUG-1992 (REL. 23, CREATED)
DT   01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   MITOCHONDRIAL RIBOSOMAL PROTEIN S19.
GN   RPS19.
OS   PETUNIA HYBRIDA (PETUNIA).
OG   MITOCHONDRION.
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC   SOLANALES; SOLANACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91252215.
RA   CONKLIN P.L., HANSON M.R.;
RL   NUCLEIC ACIDS RES. 19:2701-2705(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3704;
RX   MEDLINE; 93306753.
RA   SUTTON C.A., CONKLIN P.L., PRUITT K.D., CALFEE A.J., COBB A.G.,
RA   HANSON M.R.;
RL   CURR. GENET. 23:472-476(1993).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- POSITIONS 39, 46, 55 AND 74 ARE MODIFIED BY RNA EDITING.
CC   -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
DR   EMBL; X67028; G14164; ALT_SEQ.
DR   EMBL; X67027; G14200; ALT_SEQ.
DR   EMBL; X57283; G13340; ALT_SEQ.
DR   PIR; S29745; R3PJ19.
DR   PROSITE; PS00323; RIBOSOMAL_S19; 1.
KW   RIBOSOMAL PROTEIN; MITOCHONDRION; RNA EDITING.
SQ   SEQUENCE   94 AA;  11222 MW;  A4BE044C CRC32;

  Query Match           51.6%;  Score 48;  DB 1;  Length 94;
  Best Local Similarity  60.0%;  Pred. No. 1.61e+01;
  Matches     6;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db     72 FAFTRKRRPS 81
          : ||| |:||
Qy      1 INFTRQRQPS 10


RESULT   32
ID   VPR_HV1MN      STANDARD;      PRT;    96 AA.
AC   P05950;
DT   01-NOV-1988 (REL. 09, CREATED)
DT   01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
DT   01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN (R ORF PROTEIN).
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (MN ISOLATE) (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88219542.
RA   GURGO C., GUO H.-G., FRANCHINI G., ALDOVINI A., COLLALTI E.,
RA   FARRELL K., WONG-STAAL F., GALLO R.C., REITZ M.S. JR.;
RL   VIROLOGY 164:531-536(1988).
CC   -!- THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS PATIENT IN 1984.
DR   EMBL; M17449; G328035; -.
DR   HIV; M17449; VPR$MN.
KW   AIDS.
SQ   SEQUENCE   96 AA;  11344 MW;  D7BD031F CRC32;

  Query Match           51.6%;  Score 48;  DB 1;  Length 96;
  Best Local Similarity  42.9%;  Pred. No. 1.61e+01;
  Matches    6;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db     81 IGIIRQRRARNGAS 94
          | : |||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   33
ID   PHCA_SYNY3     STANDARD;      PRT;   162 AA.
AC   Q54715;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   C-PHYCOCYANIN ALPHA CHAIN.
GN   CPCA OR SLL1578.
OS   SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC   PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC   CYANOBACTERIA (BLUE-GREEN ALGAE); CHROOCOCCALES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96074307.
RA   PLANK T., TOOLE C., ANDERSON L.K.;
RL   J. BACTERIOL. 177:6798-6803(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97061201.
RA   KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA   MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA   HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA   SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA   TABATA S.;
RL   DNA RES. 3:109-136(1996).
RN   [3]
RP   SEQUENCE OF 1-19.
RX   MEDLINE; 97443974.
RA   SAZUKA T., OHARA O.;
RL   ELECTROPHORESIS 18:1252-1258(1997).
CC   -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC       FROM THE PHYCOBILIPROTEIN COMPLEX.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC   -!- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
DR   EMBL; U34930; G1008533; -.
DR   EMBL; D90904; G1652308; -.
KW   PHYCOBILISOME; ELECTRON TRANSPORT; PHOTOSYNTHESIS; BILE PIGMENT.
FT   BINDING      84     84       PHYCOCYANOBILIN CHROMOPHORE.
SQ   SEQUENCE   162 AA;  17586 MW;  8B79A803 CRC32;

  Query Match           51.6%;  Score 48;  DB 1;  Length 162;
  Best Local Similarity  50.0%;  Pred. No. 1.61e+01;
  Matches    6;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db     26 IAFGRLRQANAG 37
          | | | ||:: |
Qy      1 INFTRQRQPSEG 12


RESULT   34
ID   US02_HSVBS     STANDARD;      PRT;   220 AA.
AC   Q08099;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   PROTEIN US2 HOMOLOG.
OS   BOVINE HERPESVIRUS TYPE 1.2 (STRAIN ST).
OC   VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94167875.
RA   LEUNG-TACK P., AUDONNET J.F., RIVIERE M.;
RL   VIROLOGY 199:409-421(1994).
CC   -!- SIMILARITY: BELONGS TO THE HERPES VIRUSES US2 FAMILY.
DR   EMBL; Z23068; G312193; -.
SQ   SEQUENCE   220 AA;  23156 MW;  624A8A07 CRC32;

  Query Match           51.6%;  Score 48;  DB 1;  Length 220;
  Best Local Similarity  55.6%;  Pred. No. 1.61e+01;
  Matches    5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     87 LARHRQPAE 95
          ::|:|||:|
Qy      3 FTRQRQPSE 11


RESULT   35
ID   MCBC_ECOLI     STANDARD;      PRT;   272 AA.
AC   P23185;
DT   01-NOV-1991 (REL. 20, CREATED)
DT   01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   MICROCIN B17 PROCESSING PROTEIN MCBC.
GN   MCBC.
OS   ESCHERICHIA COLI.
OG   PLASMID PMCCB17.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89123111.
RA   GENILLOUD O., MORENO F., KOLTER R.;
RL   J. BACTERIOL. 171:1126-1135(1989).
CC   -!- FUNCTION: NECESSARY TO PROCESS THE INACTIVE MICROCIN B17 (MCBA)
CC       PRECURSOR INTO THE ACTIVE PEPTIDE.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC   -!- SIMILARITY: TO R.LEGUMINOSARUM TFXB WHICH IS INVOLVED IN THE
CC       PROCESSING OF TRIFOLITOXIN.
DR   EMBL; M24253; G522292; -.
DR   PIR; C32058; C32058.
KW   ANTIBIOTIC; PLASMID.
SQ   SEQUENCE   272 AA;  30753 MW;  E33FD69E CRC32;

  Query Match           51.6%;  Score 48;  DB 1;  Length 272;
  Best Local Similarity  66.7%;  Pred. No. 1.61e+01;
  Matches    6;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db     77 NFSRERLPS 85
          ||:|:| ||
Qy      2 NFTRQRQPS 10


Search completed: Tue Apr  7 08:37:41 1998
Job time : 12 secs.
```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:38:00 1998; MasPar time 6.41 Seconds
91.928 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-2
Description: (1-14) from 5541104.pep
Perfect Score: 93
Sequence: 1 INFTRQRQPSEGSS 14

Scoring table: PAM 150
Gap 15

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: sptrembl5
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phage 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 24.243; Variance 28.092; scale 0.863

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	61	65.6	188	9	Q51968	HYPOTHETICAL 20.8 KD P	3.11e-02
2	60	64.5	184	9	Q44197	ORF14.	5.36e-02
3	58	62.4	651	4	O18835	BETA-GLUCURONIDASE (EC	1.56e-01
4	57	61.3	317	2	Q14798	MAGE-4 PROTEIN.	2.65e-01
5	57	61.3	512	3	P92039	ACETYLCHOLINE RECEPTOR	2.65e-01
6	56	60.2	96	11	Q79672	VPR PROTEIN.	4.46e-01
7	56	60.2	648	10	Q61601	BETA-GLUCURONIDASE STR	4.46e-01
8	56	60.2	648	10	Q64473	BETA-GLUCORONIDASE PRE	4.46e-01
9	56	60.2	648	10	Q64474	BETA-GLUCORONIDASE PRE	4.46e-01
10	55	59.1	96	11	Q79254	VPR PROTEIN.	7.47e-01
11	55	59.1	311	3	Q23096	W01A8.5.	7.47e-01
12	55	59.1	650	1	Q01277	SULFUR CONTROLLER-2.	7.47e-01
13	54	58.1	186	8	Q40394	NGORF14 PROTEIN.	1.24e+00
14	54	58.1	373	9	O26544	SENSORY TRANSDUCTION`H	1.24e+00
15	53	57.0	96	11	O36206	VPR PROTEIN.	2.06e+00
16	53	57.0	96	11	O42054	VPR PROTEIN.	2.06e+00
17	53	57.0	96	11	Q79246	VPR PROTEIN.	2.06e+00
18	53	57.0	96	11	Q79284	VPR PROTEIN.	2.06e+00
19	53	57.0	96	11	O36203	VPR PROTEIN.	2.06e+00
20	53	57.0	96	11	O36205	VPR PROTEIN.	2.06e+00
21	53	57.0	96	11	Q75757	VPR PROTEIN.	2.06e+00
22	53	57.0	96	11	Q79282	VPR PROTEIN.	2.06e+00
23	53	57.0	96	11	Q89516	VPR PROTEIN.	2.06e+00
24	53	57.0	96	11	Q79234	VPR PROTEIN.	2.06e+00
25	53	57.0	96	11	Q79247	VPR PROTEIN.	2.06e+00
26	53	57.0	96	11	Q79244	VPR PROTEIN.	2.06e+00
27	53	57.0	96	11	Q74809	VPR PROTEIN.	2.06e+00
28	53	57.0	167	9	Q55235	FOUR ORFS, THREE COMPL	2.06e+00
29	53	57.0	270	2	Q14493	HISTONE RNA HAIRPIN-BI	2.06e+00
30	53	57.0	271	9	O28729	HYPOTHETICAL 31.0 KD P	2.06e+00
31	53	57.0	275	10	P97440	STEM-LOOP BINDING PROT	2.06e+00
32	53	57.0	1106	3	O16310	C05C8.4 PROTEIN.	2.06e+00
33	52	55.9	96	11	Q71815	VPR PROTEIN.	3.38e+00
34	52	55.9	384	11	Q69357	HOMOLOGUE OF HSV-1 GI.	3.38e+00
35	52	55.9	384	11	Q66931	GI GENE PRECURSOR.	3.38e+00
36	52	55.9	572	10	O09014	PEPTIDE/HISTIDINE TRAN	3.38e+00
37	51	54.8	596	9	O33011	HYPOTHETICAL 65.2 KD P	5.52e+00
38	50	53.8	33	11	Q74746	VPR PROTEIN (FRAGMENT)	8.94e+00
39	50	53.8	86	2	Q13447	NUP98-HOXA9 FUSION PRO	8.94e+00
40	50	53.8	174	1	O14185	HYPOTHETICAL 20.1 KD P	8.94e+00
41	50	53.8	271	2	Q99820	HOMEODOMAIN PROTEIN HO	8.94e+00
42	50	53.8	307	1	O14264	HYPOTHETICAL 35.9 KD P	8.94e+00
43	50	53.8	317	9	Q57181	PERIPLASMIC IRON-BINDI	8.94e+00
44	50	53.8	317	9	Q50988	PERIPLASMIC IRON-BINDI	8.94e+00
45	50	53.8	317	9	Q51279	PERIPLASMIC IRON-BINDI	8.94e+00
46	50	53.8	331	9	Q50964	SIMILAR TO TRANSFERRIN	8.94e+00
47	50	53.8	410	2	Q14186	E2F-RELATED TRANSCRIPT	8.94e+00
48	50	53.8	862	8	Q41594	TAXADIENE SYNTHASE.	8.94e+00
49	49	52.7	29	11	P89816	VPR PROTEIN (FRAGMENT)	1.44e+01
50	49	52.7	29	11	P89821	VPR PROTEIN (FRAGMENT)	1.44e+01
51	49	52.7	96	11	Q72500	VPR PROTEIN.	1.44e+01
52	49	52.7	343	9	P72959	HYPOTHETICAL 38.0 KD P	1.44e+01
53	49	52.7	443	11	Q98420	ASP-RICH.	1.44e+01
54	49	52.7	536	8	Q40475	BIOTIN CARBOXYLASE SUB	1.44e+01
55	49	52.7	561	2	O15373	SH3 BINDING PROTEIN 3B	1.44e+01
56	49	52.7	561	2	O00500	SH3 BINDING PROTEIN 2.	1.44e+01
57	49	52.7	561	2	P78314	SH3 BINDING PROTEIN.	1.44e+01
58	49	52.7	689	3	P90839	F16A11.1.	1.44e+01
59	49	52.7	733	10	Q62811	OXIDOSQUALENE CYCLASE.	1.44e+01
60	49	52.7	933	3	Q27772	METHYLENETETRAHYDROFOL	1.44e+01
61	48	51.6	96	11	O36202	VPR PROTEIN.	2.30e+01
62	48	51.6	96	11	O42084	VPR PROTEIN.	2.30e+01
63	48	51.6	96	11	O36207	VPR PROTEIN.	2.30e+01
64	48	51.6	96	11	P88434	VPR PROTEIN.	2.30e+01
65	48	51.6	96	11	Q73332	C18MBC, COMPLETE SEQUE	2.30e+01
66	48	51.6	96	11	Q71944	CNTRL 2, CLONE 19, ACC	2.30e+01
67	48	51.6	96	11	Q79249	VPR PROTEIN.	2.30e+01
68	48	51.6	96	11	Q71938	CNTRL 2, CLONE 18, ACC	2.30e+01
69	48	51.6	96	11	O36204	VPR PROTEIN.	2.30e+01
70	48	51.6	96	11	Q79242	VPR PROTEIN.	2.30e+01
71	48	51.6	96	11	Q79793	ISOLATE NY5, COMPLETE	2.30e+01
72	48	51.6	142	3	Q21371	K08H2.7.	2.30e+01
73	48	51.6	165	9	O25152	PURINE-BINDING CHEMOTA	2.30e+01
74	48	51.6	220	11	O39502	US2 PROTEIN.	2.30e+01
75	48	51.6	521	3	Q25429	EXTRACELLULAR MATRIX P	2.30e+01
76	48	51.6	942	11	Q65729	VIRAL POLYMERASE, COAT	2.30e+01
77	48	51.6	1235	3	Q24315	TYROSINE KINASE (FRAGM	2.30e+01
78	48	51.6	3093	11	Q65730	POLYPROTEIN.	2.30e+01
79	47	50.5	111	2	Q92528	5'-TERMINAL REGION OF	3.64e+01
80	47	50.5	198	1	Q06898	DNA FOR ORFS AND PPR1	3.64e+01
81	47	50.5	199	9	O34723	SENSOR REGULATOR.	3.64e+01
82	47	50.5	262	12	Q90676	ROUS SARCOMA VIRUS TRA	3.64e+01
83	47	50.5	287	12	Q90376	Y-BOX BINDING PROTEIN.	3.64e+01
84	47	50.5	288	9	O25668	CO-CHAPERONE-CURVED DN	3.64e+01
85	47	50.5	297	9	O31822	YNGB PROTEIN.	3.64e+01
86	47	50.5	322	10	Q60951	MYB-1B.	3.64e+01
87	47	50.5	324	4	Q28618	MRNP PARTICLE MAJOR PR	3.64e+01
88	47	50.5	326	12	Q91956	SIMILAR TO DBPB/YB-1 O	3.64e+01
89	47	50.5	348	12	Q08243	CCAAT-BINDING PROTEIN	3.64e+01
90	47	50.5	370	11	Q86789	GLYCOPROTEIN GI.	3.64e+01
91	47	50.5	397	1	Q07949	CHROMOSOME XII READING	3.64e+01
92	47	50.5	488	1	O14011	BETA-TRANSDUCIN.	3.64e+01
93	47	50.5	527	9	P71590	HYPOTHETICAL 56.9 KD P	3.64e+01

94	47	50.5	658	3	Q09946	HYPOTHETICAL 75.1 KD P	3.64e+01
95	47	50.5	958	8	Q40554	PNLA-35.	3.64e+01
96	47	50.5	977	3	Q94678	VALS GENE (FRAGMENT).	3.64e+01
97	47	50.5	1048	9	O05884	HYPOTHETICAL 110.2 KD	3.64e+01
98	47	50.5	1142	11	P89462	RIBONUCLEOTIDE REDUCTA	3.64e+01
99	47	50.5	1180	2	Q92625	MYELOBLAST KIAA0229 (F	3.64e+01
100	47	50.5	1937	9	O30482	PKS MODULE 4.	3.64e+01

ALIGNMENTS

```
RESULT    1
ID   Q51968          PRELIMINARY;   PRT;   188 AA.
AC   Q51968;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 20.8 KD PROTEIN.
OS   AGROBACTERIUM RHIZOGENES.
OG   PLASMID PRI1724.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   RHIZOBIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF03-01724;
RX   MEDLINE; 94227335.
RA   TANAKA N., IKEDA T., OKA A.;
RL   BIOSCI. BIOTECHNOL. BIOCHEM. 58:548-551(1994).
RN   [2]
RP   REVISIONS.
RC   STRAIN=MAFF03-01724;
RA   TANAKA N.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AB006689; D1023202; -.
KW   HYPOTHETICAL PROTEIN; PLASMID.
SQ   SEQUENCE   188 AA;  20769 MW;  7FF52496 CRC32;

  Query Match           65.6%;  Score 61;  DB 9;  Length 188;
  Best Local Similarity  72.7%;  Pred. No. 3.11e-02;
  Matches     8;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    161 FTRQRQPEDQS 171
          ||||||| : |
Qy      3 FTRQRQPSEGS 13


RESULT    2
ID   Q44197          PRELIMINARY;   PRT;   184 AA.
AC   Q44197;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   ORF14.
OS   AGROBACTERIUM RHIZOGENES.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   RHIZOBIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91352070.
RA   HANSEN G., LARRIBE M., VAUBERT D., TEMPE J., BIERMANN B.J.,
RA   MONTOYA A.L., CHILTON M.D., BREVET J.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 88:7763-7767(1991).
DR   EMBL; M60490; G142251; -.
SQ   SEQUENCE   184 AA;  20254 MW;  73F2711C CRC32;

  Query Match           64.5%;  Score 60;  DB 9;  Length 184;
  Best Local Similarity  72.7%;  Pred. No. 5.36e-02;
  Matches     8;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    157 FTRQRQPQDQS 167
          ||||||| : |
Qy      3 FTRQRQPSEGS 13


RESULT    3
ID   O18835          PRELIMINARY;   PRT;   651 AA.
AC   O18835;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   BETA-GLUCURONIDASE (EC 3.2.1.31).
GN   GUSB.
OS   CANIS FAMILIARIS (DOG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; CARNIVORA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   RAY J., BOUVET A.B., DESANTO C., FYFE J.C., WU D., WOLFE J.H.,
RA   AGUIRRE G.D., PATTERSON D.F., HASKINS M.E., HENTHORN P.S.;
RL   SUBMITTED (SEP-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF019759; G2425091; -.
DR   PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
DR   PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
KW   HYDROLASE; GLYCOSIDASE.
SQ   SEQUENCE   651 AA;  74433 MW;  6CA3E409 CRC32;

  Query Match           62.4%;  Score 58;  DB 4;  Length 651;
  Best Local Similarity  63.6%;  Pred. No. 1.56e-01;
  Matches     7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    608 FTRQRQPKAAA 618
          |||||||  ::
Qy      3 FTRQRQPSEGS 13


RESULT    4
ID   Q14798          PRELIMINARY;   PRT;   317 AA.
AC   Q14798;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   MAGE-4 PROTEIN.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95369706.
RA   IMAI Y., SHICHIJO S., YAMADA A., KATAYAMA T., YANO H., ITOH K.;
RL   GENE 160:287-290(1995).
DR   EMBL; D32075; G1125014; -.
SQ   SEQUENCE   317 AA;  35044 MW;  982B1AC9 CRC32;

  Query Match           61.3%;  Score 57;  DB 2;  Length 317;
  Best Local Similarity  71.4%;  Pred. No. 2.65e-01;
  Matches    10;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db     76 ISFTCWRQPNEGSS 89
          |:||  |||:||||
Qy      1 INFTRQRQPSEGSS 14


RESULT    5
ID   P92039          PRELIMINARY;   PRT;   512 AA.
AC   P92039;
DT   01-MAY-1997 (TREMBLREL. 03, CREATED)
DT   01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   ACETYLCHOLINE RECEPTOR ALPHA SUBUNIT PRECURSOR.
GN   ACHR.
OS   HAEMONCHUS CONTORTUS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; SPIRURIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97237560.
RA   HOEKSTRA R., VISSER A., WILEY L., WEISS A., SANGSTER N., ROOS M.;
```

```
RL   MOL. BIOCHEM. PARASITOL. 84:179-187(1997).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR   EMBL; U72490; G1894921; -.
DR   PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW   SIGNAL; POSTSYNAPTIC MEMBRANE; IONIC CHANNEL; GLYCOPROTEIN;
KW   TRANSMEMBRANE.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    512       ACETYLCHOLINE RECEPTOR ALPHA SUBUNIT.
SQ   SEQUENCE   512 AA;  59540 MW;  515C7A4A CRC32;

  Query Match           61.3%;  Score 57;  DB 3;  Length 512;
  Best Local Similarity  38.5%;  Pred. No. 2.65e-01;
  Matches     5;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db     32 VNYNRHRRPASGP 44
          :|: |:|:|: |:
Qy      1 INFTRQRQPSEGS 13


RESULT    6
ID   Q79672         PRELIMINARY;   PRT;    96 AA.
AC   Q79672;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90034200.
RA   SPIRE B., SIRE J., ZACHAR V., REY F., BARRE-SINOUSSI F.,
RA   GALIBERT F., HAMPE A., CHERMANN J.C.;
RL   GENE 81:275-284(1989).
DR   EMBL; M27323; G328160; -.
SQ   SEQUENCE   96 AA;  11384 MW;  D187EEAA CRC32;

  Query Match           60.2%;  Score 56;  DB 11;  Length 96;
  Best Local Similarity  57.1%;  Pred. No. 4.46e-01;
  Matches     8;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db     81 ISITRQRRARNGSS 94
          |::||||:: :|||
Qy      1 INFTRQRQPSEGSS 14


RESULT    7
ID   Q61601         PRELIMINARY;   PRT;   648 AA.
AC   Q61601;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   BETA-GLUCURONIDASE STRUCTURAL PRECURSOR (GUS-S-A ALLELE).
GN   GUS-S.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88216590.
RA   FUNKENSTEIN B., LEARY S.L., STEIN J.C., CATTERALL J.F.;
RL   MOL. CELL. BIOL. 8:1160-1168(1988).
DR   EMBL; M19279; G309257; -.
DR   MGD; MGI:95874; GUS-S.
DR   PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
DR   PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
KW   SIGNAL.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    648       POTENTIAL.
SQ   SEQUENCE   648 AA;  74192 MW;  567702FB CRC32;

  Query Match           60.2%;  Score 56;  DB 10;  Length 648;
  Best Local Similarity  100.0%;  Pred. No. 4.46e-01;
  Matches     7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    605 FTRQRQP 611
          |||||||
Qy      3 FTRQRQP 9


RESULT    8
ID   Q64473         PRELIMINARY;   PRT;   648 AA.
AC   Q64473;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   BETA-GLUCORONIDASE PRECURSOR (EC 3.2.1.31) (BETA-GLUCURONIDASE).
GN   GUS-S.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YBR;
RX   MEDLINE; 89384641.
RA   WAWRZYNIAK C.J., GALLAGHER P.M., D'AMORE M.A., CARTER J.E., LUND S.D.,
RA   RINCHIK E.M., GANSCHOW R.E.;
RL   MOL. CELL. BIOL. 9:4074-4078(1989).
CC   -!- CATALYTIC ACTIVITY: A BETA-D-GLUCURONOSIDE + H(2)O = AN ALCOHOL +
CC       D-GLUCURONATE.
DR   EMBL; M28540; G193719; -.
DR   PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
DR   PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
KW   SIGNAL; HYDROLASE; GLYCOSIDASE.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    648       BETA-GLUCORONIDASE.
SQ   SEQUENCE   648 AA;  74195 MW;  1E21D33B CRC32;

  Query Match           60.2%;  Score 56;  DB 10;  Length 648;
  Best Local Similarity  100.0%;  Pred. No. 4.46e-01;
  Matches     7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    605 FTRQRQP 611
          |||||||
Qy      3 FTRQRQP 9


RESULT    9
ID   Q64474         PRELIMINARY;   PRT;   648 AA.
AC   Q64474;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   BETA-GLUCORONIDASE PRECURSOR (EC 3.2.1.31) (BETA-GLUCURONIDASE).
GN   GUS-S.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C3H/HEJ; TISSUE=SPERM;
RX   MEDLINE; 89384641.
RA   WAWRZYNIAK C.J., GALLAGHER P.M., D'AMORE M.A., CARTER J.E., LUND S.D.,
RA   RINCHIK E.M., GANSCHOW R.E.;
RL   MOL. CELL. BIOL. 9:4074-4078(1989).
CC   -!- CATALYTIC ACTIVITY: A BETA-D-GLUCURONOSIDE + H(2)O = AN ALCOHOL +
CC       D-GLUCURONATE.
DR   EMBL; M28541; G193721; -.
DR   PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
DR   PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
KW   SIGNAL; HYDROLASE; GLYCOSIDASE.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    648       BETA-GLUCORONIDASE.
SQ   SEQUENCE   648 AA;  74207 MW;  408B6BC7 CRC32;
```

```
  Query Match           60.2%;  Score 56;  DB 10;  Length 648;
  Best Local Similarity 100.0%;  Pred. No. 4.46e-01;
  Matches    7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    605 FTRQRQP 611
          |||||||
Qy      3 FTRQRQP 9


RESULT   10
ID   Q79254         PRELIMINARY;   PRT;    96 AA.
AC   Q79254;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT 3499, HOMOSEXUAL, GERMAN;
RA   KUIKEN L., CORNELISSEN E., ZORGDRAGER F., HARTMAN S., GIBBS J.,
RA   GOUDSMIT J.;
RL   J. GEN. VIROL. 0:0-0(1996).
DR   EMBL; Z68558; E218342; -.
SQ   SEQUENCE   96 AA;  11252 MW;  48376E02 CRC32;

  Query Match           59.1%;  Score 55;  DB 11;  Length 96;
  Best Local Similarity 57.1%;  Pred. No. 7.47e-01;
  Matches    8;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGSS 94
          | :||||:: :|||
Qy      1 INFTRQRQPSEGSS 14


RESULT   11
ID   Q23096         PRELIMINARY;   PRT;   311 AA.
AC   Q23096;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   W01A8.5.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   WILKINSON J.;
RL   SUBMITTED (APR-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,
RA   COULSON A., CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A.,
RA   FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
RA   JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
RA   LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B.,
RA   O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A.,
RA   SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A., SONNHAMMER E.,
RA   STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M.,
RA   VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA   WILKINSON-SPROAT J., WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
DR   EMBL; Z71267; E237025; -.
SQ   SEQUENCE   311 AA;  35099 MW;  CDF5715F CRC32;

  Query Match           59.1%;  Score 55;  DB 3;  Length 311;
  Best Local Similarity 57.1%;  Pred. No. 7.47e-01;
  Matches    8;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    298 INFTKDNQPSTSSN 311
          ||||:: ||| :|:
Qy      1 INFTRQRQPSEGSS 14


RESULT   12
ID   Q01277         PRELIMINARY;   PRT;   650 AA.
AC   Q01277;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   SULFUR CONTROLLER-2.
GN   SCON-2.
OS   NEUROSPORA CRASSA.
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; PYRENOMYCETES; SORDARIALES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WILD-TYPE 74-OR23-1A;
RX   MEDLINE; 95241499.
RA   KUMAR A., PAIETTA J.V.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 92:3343-3347(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WILD-TYPE 74-OR23-1A;
RX   MEDLINE; 90377210.
RA   PAIETTA J.V.;
RL   MOL. CELL. BIOL. 10:5207-5214(1990).
DR   EMBL; U17251; G806758; -.
DR   PROSITE; PS00678; G_BETA_REPEATS; 2.
KW   REPEAT.
SQ   SEQUENCE   650 AA;  72189 MW;  59E1F24E CRC32;

  Query Match           59.1%;  Score 55;  DB 1;  Length 650;
  Best Local Similarity 53.8%;  Pred. No. 7.47e-01;
  Matches    7;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    196 NYTRQRQLAKGGP 208
          |:|||| : |::
Qy      2 NFTRQRQPSEGSS 14


RESULT   13
ID   Q40394         PRELIMINARY;   PRT;   186 AA.
AC   Q40394;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   NGORF14 PROTEIN.
GN   NGORF14.
OS   NICOTIANA GLAUCA (GLAUCOUS TOBACCO) (TREE TOBACCO).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC   SOLANALES; SOLANACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94301302.
RA   AOKI S., KAWAOKA A., SEKINE M., ICHIKAWA T., FUJITA T.,
RA   SHINMYO A., SYONO K.;
RL   MOL. GEN. GENET. 243:706-710(1994).
DR   EMBL; D16559; G1113089; -.
SQ   SEQUENCE   186 AA;  20952 MW;  540A964C CRC32;

  Query Match           58.1%;  Score 54;  DB 8;  Length 186;
  Best Local Similarity 63.6%;  Pred. No. 1.24e+00;
  Matches    7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    159 FTRQHQPQDPS 169
          ||||:|| : |
Qy      3 FTRQRQPSEGS 13


RESULT   14
ID   O26544         PRELIMINARY;   PRT;   373 AA.
```

```
AC   O26544;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   SENSORY TRANSDUCTION HISTIDINE KINASE.
GN   MTH444.
OS   METHANOBACTERIUM THERMOAUTOTROPHICUM.
OC   ARCHAEBACTERIA; EURYARCHAEOTA; METHANOBACTERIALES;
OC   METHANOBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DELTA H;
RA   SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
RA   ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
RA   HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,
RA   SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R., JIWANI N.,
RA   CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S., MCDOUGALL S.,
RA   SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M., DANIELS C.J.,
RA   MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
RL   J. BACTERIOL. 179:7135-7155(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DELTA H;
RA   SMITH D.R.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AE000828; G2621510; -.
SQ   SEQUENCE   373 AA;  42263 MW;  A636962D CRC32;

  Query Match           58.1%;  Score 54;  DB 9;  Length 373;
  Best Local Similarity  60.0%;  Pred. No. 1.24e+00;
  Matches     6;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    277 IKFTRDRDPA 286
          |:|||:|:|:
Qy      1 INFTRQRQPS 10


RESULT   15
ID   O36206         PRELIMINARY;   PRT;    96 AA.
AC   O36206;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   SONG J., WANG B., GE Y.C., DWYER D., DOWTON D., CUNNINGHAM A.,
RA   SAKSENA N.;
RL   SUBMITTED (SEP-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF000318; G2393859; -.
SQ   SEQUENCE   96 AA;  11323 MW;  3AFB9942 CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   16
ID   O42054         PRELIMINARY;   PRT;    96 AA.
AC   O42054;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   SONG J., WANG B., GE Y.C., DWYER D., DOWTON D., CUNNINGHAM A.,
RA   SAKSENA N.;
RL   SUBMITTED (SEP-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF000321; G2393865; -.
DR   EMBL; AF000320; G2393863; -.
SQ   SEQUENCE   96 AA;  11351 MW;  66A4027D CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   17
ID   Q79246         PRELIMINARY;   PRT;    96 AA.
AC   Q79246;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT 0709, HOMOSEXUAL, DUTCH;
RA   KUIKEN L., CORNELISSEN E., ZORGDRAGER F., HARTMAN S., GIBBS J.,
RA   GOUDSMIT J.;
RL   J. GEN. VIROL. 0:0-0(1996).
DR   EMBL; Z68551; E218426; -.
SQ   SEQUENCE   96 AA;  11488 MW;  21D0C46C CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   18
ID   Q79284         PRELIMINARY;   PRT;    96 AA.
AC   Q79284;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT 0070, IV DRUG USER, DUTCH;
RA   KUIKEN L., CORNELISSEN E., ZORGDRAGER F., HARTMAN S., GIBBS J.,
RA   GOUDSMIT J.;
RL   J. GEN. VIROL. 0:0-0(1996).
DR   EMBL; Z68585; E218358; -.
SQ   SEQUENCE   96 AA;  11327 MW;  E04768CA CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;
```

```
Db      81 IGITRQRRARNGAS 94
           | :||||:: :|:|
Qy       1 INFTRQRQPSEGSS 14


RESULT   19
ID   O36203         PRELIMINARY;   PRT;    96 AA.
AC   O36203;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   SONG J., WANG B., GE Y.C., DWYER D., DOWTON D., CUNNINGHAM A.,
RA   SAKSENA N.;
RL   SUBMITTED (SEP-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF000315; G2393853; -.
SQ   SEQUENCE   96 AA;  11349 MW;  C0013AFA CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches    7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db      81 IGITRQRRARNGAS 94
           | :||||:: :|:|
Qy       1 INFTRQRQPSEGSS 14


RESULT   20
ID   O36205         PRELIMINARY;   PRT;    96 AA.
AC   O36205;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   SONG J., WANG B., GE Y.C., DWYER D., DOWTON D., CUNNINGHAM A.,
RA   SAKSENA N.;
RL   SUBMITTED (SEP-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF000317; G2393857; -.
SQ   SEQUENCE   96 AA;  11367 MW;  0840C3DA CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches    7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db      81 IGITRQRRARNGAS 94
           | :||||:: :|:|
Qy       1 INFTRQRQPSEGSS 14


RESULT   21
ID   Q75757         PRELIMINARY;   PRT;    96 AA.
AC   Q75757;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JRFL; TISSUE=BRAIN;
RA   KOYANAGI Y., MILES S., MITSUYASU R.T., MERRILL J.E., VINTERS H.V.,
RA   CHEN I.S.;
RL   SCIENCE 236:819-822(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JRFL; TISSUE=BRAIN;
RX   MEDLINE; 91043044.
RA   O'BRIEN W.A., KOYANAGI Y., NAMAZIE A., ZHAO J.Q., DIAGNE A.,
RA   IDLER K., ZACK J.A., CHEN I.S.;
RL   NATURE 348:69-73(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JRFL; TISSUE=BRAIN;
RX   MEDLINE; 92092169.
RA   PANG S., VINTERS H.V., AKASHI T., O'BRIEN W.A., CHEN I.S.;
RL   J. ACQUIR. IMMUNE DEFIC. SYNDR. 4:1082-1092(1991).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JRFL; TISSUE=BRAIN;
RA   PANG S., VINTERS H.V., AKASHI T., O'BRIEN W.A., CHEN I.S.,
RA   KOYANAGI Y., NAMAZIE A., ZHAO J., DIAGNE A., IDLER K.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U63632; G1465784; -.
SQ   SEQUENCE   96 AA;  11377 MW;  50CD3483 CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches    7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db      81 IGITRQRRARNGAS 94
           | :||||:: :|:|
Qy       1 INFTRQRQPSEGSS 14


RESULT   22
ID   Q79282         PRELIMINARY;   PRT;    96 AA.
AC   Q79282;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT 3057, IV DRUG USER, DUTCH;
RA   KUIKEN L., CORNELISSEN E., ZORGDRAGER F., HARTMAN S., GIBBS J.,
RA   GOUDSMIT J.;
RL   J. GEN. VIROL. 0:0-0(1996).
DR   EMBL; Z68584; E218438; -.
SQ   SEQUENCE   96 AA;  11369 MW;  4C351F74 CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches    7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db      81 IGITRQRRARNGAS 94
           | :||||:: :|:|
Qy       1 INFTRQRQPSEGSS 14


RESULT   23
ID   Q89516         PRELIMINARY;   PRT;    96 AA.
AC   Q89516;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
```

```
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT LW;
RA   WANG B., CHUN GE Y.C., PALASANTHIRAN P., XIANG S.H., ZIEGLER J.,
RA   RANDLE C., DOWTON D., CUNNINGHAM A., SAKSENA N.;
RL   SUBMITTED (AUG-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT LW;
RA   BLOUIN C.;
RL   SUBMITTED (JUN-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
SQ   SEQUENCE   96 AA;  11323 MW;  36B8DAC9 CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   24
ID   Q79234         PRELIMINARY;   PRT;    96 AA.
AC   Q79234;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT 0140, HOMOSEXUAL, DUTCH;
RA   KUIKEN L., CORNELISSEN E., ZORGDRAGER F., HARTMAN S., GIBBS J.,
RA   GOUDSMIT J.;
RL   J. GEN. VIROL. 0:0-0(1996).
DR   EMBL; Z68540; E218330; -.
SQ   SEQUENCE   96 AA;  11338 MW;  E5898B92 CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   25
ID   Q79247         PRELIMINARY;   PRT;    96 AA.
AC   Q79247;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT 1145, HOMOSEXUAL, DUTCH;
RA   KUIKEN L., CORNELISSEN E., ZORGDRAGER F., HARTMAN S., GIBBS J.,
RA   GOUDSMIT J.;
RL   J. GEN. VIROL. 0:0-0(1996).
DR   EMBL; Z68552; E218337; -.
SQ   SEQUENCE   96 AA;  11323 MW;  250A3402 CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   26
ID   Q79244         PRELIMINARY;   PRT;    96 AA.
AC   Q79244;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT 0617, HOMOSEXUAL, DUTCH;
RA   KUIKEN L., CORNELISSEN E., ZORGDRAGER F., HARTMAN S., GIBBS J.,
RA   GOUDSMIT J.;
RL   J. GEN. VIROL. 0:0-0(1996).
DR   EMBL; Z68549; E218425; -.
SQ   SEQUENCE   96 AA;  11378 MW;  D20333D8 CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGTS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   27
ID   Q74809         PRELIMINARY;   PRT;    96 AA.
AC   Q74809;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D31;
RX   MEDLINE; 93090480.
RA   KREUTZ R., DIETRICH U., KUHNEL H., NIESELT-STRUWE K., EIGEN M.,
RA   RUBSAMEN-WAIGMANN H.;
RL   AIDS RES. HUM. RETROVIRUSES 8:1619-1629(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D31;
RA   BRYANT B.;
RL   SUBMITTED (DEC-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U43096; G1171168; -.
SQ   SEQUENCE   96 AA;  11386 MW;  C2597A11 CRC32;

  Query Match           57.0%;  Score 53;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     81 IGITRQRRARNGAS 94
          | :||||:: :|:|
Qy      1 INFTRQRQPSEGSS 14
```

```
RESULT   28
ID   Q55235         PRELIMINARY;   PRT;   167 AA.
AC   Q55235;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   FOUR ORFS, THREE COMPLETE, AND ONE 3' END.
OS   SYNECHOCOCCUS SP.
OC   PROKARYOTA; BACTERIA; GRACILICUTES; OXYPHOTOBACTERIA;
OC   CYANOBACTERIA; CHROOCOCCALES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94064558.
RA   DOLGANOV N., GROSSMAN A.R.;
RL   J. BACTERIOL. 175:7644-7651(1993).
DR   EMBL; L19521; G310859; -.
SQ   SEQUENCE   167 AA;  18740 MW;  E26339DB CRC32;

  Query Match           57.0%;  Score 53;  DB 9;  Length 167;
  Best Local Similarity  50.0%;  Pred. No. 2.06e+00;
  Matches     5;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db     74 VNYARQRQAG 83
          :|::||||:: .
Qy      1 INFTRQRQPS 10


RESULT   29
ID   Q14493         PRELIMINARY;   PRT;   270 AA.
AC   Q14493;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   HISTONE RNA HAIRPIN-BINDING PROTEIN.
GN   SLBP.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97201520.
RA   MARTIN F., SCHALLER A., EGLITE S., SCHUEMPERLI D., MUELLER B.;
RL   EMBO J. 16:769-778(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93251024.
RA   MCCOMBIE W.R., MARTIN-GALLARDO A., GOCAYNE J.D., FITZGERALD M.,
RA   DUBNICK M., KELLEY J.M., CASTILLA L., LIU L.I., WALLACE S., TRAPP S.,
RA   TAGLE D., WHALEY W.L., CHENG S., GUSELLA J., FRISCHAUF A.M.,
RA   POUSTKA A., LEHRACH H., COLLINS F.S., KERLAVAGE A.R., FIELDS C.,
RA   VENTER J.C.;
RL   NATURE GENET. 1:348-353(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   WANG Z.F., WHITFIELD M.L., INGLEDUE I.T.C., III DOMINSKI Z.,
RA   MARZLUFF W.F.;
RL   SUBMITTED (DEC-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z71188; E236578; -.
DR   EMBL; U75679; G1732077; -.
SQ   SEQUENCE   270 AA;  31285 MW;  099935BE CRC32;

  Query Match           57.0%;  Score 53;  DB 2;  Length 270;
  Best Local Similarity  53.8%;  Pred. No. 2.06e+00;
  Matches     7;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db    103 DFGRERKSSSGSS 115
          :| |:| :| |||
Qy      2 NFTRQRQPSEGSS 14


RESULT  30
```

```
ID   O28729         PRELIMINARY;   PRT;   271 AA.
AC   O28729;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 31.0 KD PROTEIN.
GN   AF1543.
OS   ARCHAEOGLOBUS FULGIDUS.
OC   ARCHAEBACTERIA; EURYARCHAEOTA; ARCHAEOGLOBALES; ARCHAEOGLOBACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   KLENK H.P., CLAYTON R.A., TOMB J., WHITE O., NELSON K.E.,
RA   KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,
RA   RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,
RA   FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
RA   KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
RA   PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,
RA   OVERBEEK R., GOCAYNE J.D., WEIDMAN J.F., MCDONALD L., UTTERBACK T.,
RA   COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,
RA   SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,
RA   MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,
RA   VENTER J.C.;
RL   SUBMITTED (DEC-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   KLENK H.P., CLAYTON R.A., TOMB J., WHITE O., NELSON K.E.,
RA   KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,
RA   RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,
RA   FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
RA   KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
RA   PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,
RA   OVERBEEK R., GOCAYNE J.D., WEIDMAN J.F., MCDONALD L., UTTERBACK T.,
RA   COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,
RA   SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,
RA   MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,
RA   VENTER J.C.;
RL   SUBMITTED (NOV-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AE000996; G2649029; -.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   271 AA;  31009 MW;  D2510E8B CRC32;

  Query Match           57.0%;  Score 53;  DB 9;  Length 271;
  Best Local Similarity  41.7%;  Pred. No. 2.06e+00;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db     11 DFVRHRQNEDGG 22
          :| |:||  :|:
Qy      2 NFTRQRQPSEGS 13


RESULT   31
ID   P97440         PRELIMINARY;   PRT;   275 AA.
AC   P97440;
DT   01-MAY-1997 (TREMBLREL. 03, CREATED)
DT   01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   STEM-LOOP BINDING PROTEIN (HISTONE STEM-LOOP BINDING PROTEIN).
GN   SLBP.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIH SWISS;
RA   WANG Z.F., WHITFIELD M.L., INGLEDUE I.T.C., III DOMINSKI Z.,
RA   MARZLUFF W.F.;
RL   SUBMITTED (DEC-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U75680; G1732079; -.
DR   MGD; MGI:108402; SLBP.
SQ   SEQUENCE   275 AA;  31603 MW;  1793AC0A CRC32;

  Query Match           57.0%;  Score 53;  DB 10;  Length 275;
  Best Local Similarity  53.8%;  Pred. No. 2.06e+00;
```

```
  Matches     7;  Conservative     3;  Mismatches   3;  Indels    0;  Gaps    0;

Db    103 DFGRERKSSSGSS 115
          :| |:| :| |||
Qy      2 NFTRQRQPSEGSS 14


RESULT   32
ID   O16310         PRELIMINARY;   PRT;  1106 AA.
AC   O16310;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   C05C8.4 PROTEIN.
GN   C05C8.4.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., BONFIELD J.,
RA   BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A., CRAXTON M.,
RA   DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L., GARDNER A., GREEN P.,
RA   HAWKINS T., HILLIER L., JIER M., JOHNSTON L., JONES M., KERSHAW J.,
RA   KIRSTEN J., LAISTER N., LATREILLE P., LIGHTNING J., LLOYD C.,
RA   MCMURRAY A., MORTIMORE B., O'CALLAGHAN M., PARSONS J., PERCY C.,
RA   RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A.,
RA   SONNHAMMER E., STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K.,
RA   VAUDIN M., VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA   WILKINSON-SPROAT J., WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   SAMMONS L., WOHLDMANN P.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   WATERSTON R.;
RL   SUBMITTED (JUL-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF016430; G2291255; -.
SQ   SEQUENCE   1106 AA;  122362 MW;  3921B0DE CRC32;

  Query Match           57.0%;  Score 53;  DB 3;  Length 1106;
  Best Local Similarity  46.2%;  Pred. No. 2.06e+00;
  Matches     6;  Conservative     5;  Mismatches   2;  Indels    0;  Gaps    0;

Db     71 LSFFKQRQPASGG 83
          ::| :|||: |:
Qy      1 INFTRQRQPSEGS 13


RESULT   33
ID   Q71815         PRELIMINARY;   PRT;    96 AA.
AC   Q71815;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   VPR PROTEIN.
GN   VPR.
OS   HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC   LENTIVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   ZHU T., HO D.D.;
RL   SUBMITTED (MAY-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   MACINNES K.A.;
RL   SUBMITTED (MAR-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U23487; G818218; -.
SQ   SEQUENCE   96 AA;  11410 MW;  A9D2B961 CRC32;

  Query Match           55.9%;  Score 52;  DB 11;  Length 96;
  Best Local Similarity  50.0%;  Pred. No. 3.38e+00;
  Matches     7;  Conservative     4;  Mismatches   3;  Indels    0;  Gaps    0;

Db     81 IGITRQRRTRNGAS 94
          | :|||:  :|:|
Qy      1 INFTRQRQPSEGSS 14


RESULT   34
ID   Q69357         PRELIMINARY;   PRT;   384 AA.
AC   Q69357;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   HOMOLOGUE OF HSV-1 GI.
OS   FELINE HERPESVIRUS (FELID HERPESVIRUS 1).
OC   VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=G2620;
RA   SONDERMEIJER P.J.A.;
RL   SUBMITTED (JUL-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=G2620;
RX   MEDLINE; 95266277.
RA   WILLEMSE M.J., STRIJDVEEN I.G.L., VAN SCHOONEVELD S.H.B.,
RA   DEN BERG M.C., SONDERMEIJER P.J.A.;
RL   VIROLOGY 208:704-711(1995).
DR   EMBL; D42113; G893371; -.
SQ   SEQUENCE   384 AA;  43009 MW;  A1BFA786 CRC32;

  Query Match           55.9%;  Score 52;  DB 11;  Length 384;
  Best Local Similarity  66.7%;  Pred. No. 3.38e+00;
  Matches     8;  Conservative     2;  Mismatches   2;  Indels    0;  Gaps    0;

Db    343 FTRQTKPSNSSS 354
          |||| ||::||
Qy      3 FTRQRQPSEGSS 14


RESULT   35
ID   Q66931         PRELIMINARY;   PRT;   384 AA.
AC   Q66931;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE   GI GENE PRECURSOR.
GN   GI.
OS   FELINE HERPESVIRUS (FELID HERPESVIRUS 1).
OC   VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B927;
RA   MIJNES J.D.F., VAN DER HORST L.M., VAN ANKEN E., HORZINEK M.C.,
RA   ROTTIER P.J.M., DE GROOT R.J.;
RL   J. VIROL. 70:5466-5475(1996).
DR   EMBL; X98448; E249342; -.
KW   SIGNAL.
FT   SIGNAL        1     18       POTENTIAL.
SQ   SEQUENCE   384 AA;  43019 MW;  66F2059A CRC32;

  Query Match           55.9%;  Score 52;  DB 11;  Length 384;
  Best Local Similarity  66.7%;  Pred. No. 3.38e+00;
  Matches     8;  Conservative     2;  Mismatches   2;  Indels    0;  Gaps    0;

Db    343 FTRQTKPSNSSS 354
```

```
          ||||  ||::||
Qy      3 FTRQRQPSEGSS 14

Search completed: Tue Apr  7 08:38:14 1998
Job time : 14 secs.
```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:41:04 1998; MasPar time 2.35 Seconds
86.413 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-3
Description: (1-12) from 5541104.pep
Perfect Score: 82
Sequence: 1 LFRAVITKKVAD 12

Scoring table: PAM 150
Gap 15

Searched: 111725 seqs, 16919825 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: a-geneseq30
1:a-geneseq1

Statistics: Mean 19.964; Variance 25.854; scale 0.772

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	82	100.0	38	1	R80619	Immunogenic peptide of	1.60e-08
2	82	100.0	335	1	R70909	Human melanoma antigen	1.60e-08
3	65	79.3	36	1	R47325	HLA-A3 MAGE 1 antigen	3.04e-04
4	58	70.7	35	1	R47324	HLA-A3 MAGE 1 antigen	1.37e-02
5	58	70.7	35	1	R65120	MAGE 1 immunogenic pep	1.37e-02
6	58	70.7	35	1	R49228	HLA-A11 MAGE 1 antigen	1.37e-02
7	58	70.7	36	1	R65125	MAGE 1 immunogenic pep	1.37e-02
8	58	70.7	36	1	R49230	HLA-A11 MAGE 1 antigen	1.37e-02
9	48	58.5	604	1	W11578	Bacillus licheniformis	2.22e+00
10	47	57.3	335	1	R67916	(1-3)-beta-D-glucan se	3.59e+00
11	47	57.3	748	1	R98227	Rat neuronal protein k	3.59e+00
12	46	56.1	166	1	P91891	Part of the sequence o	5.77e+00
13	46	56.1	356	1	P81996	Sequence encoded by no	5.77e+00
14	46	56.1	569	1	W07702	Mouse ETS2 repressor f	5.77e+00
15	46	56.1	574	1	W07700	Human ETS2 repressor f	5.77e+00
16	46	56.1	754	1	R90617	Sulfolobus solfataricu	5.77e+00
17	45	54.9	185	1	R60900	Borrelia VSDA antigen	9.22e+00
18	45	54.9	199	1	R62793	Borrelia KL11 antigen	9.22e+00
19	45	54.9	199	1	R60908	Borrelia PBI antigen v	9.22e+00
20	45	54.9	200	1	R62786	Borrelia VSDA antigen	9.22e+00
21	45	54.9	201	1	R62784	Borrelia M57 antigen v	9.22e+00
22	45	54.9	203	1	R60904	Borrelia IP90 antigen	9.22e+00
23	45	54.9	203	1	R62792	Borrelia BITS antigen	9.22e+00
24	45	54.9	203	1	R62790	Borrelia IP90 antigen	9.22e+00
25	45	54.9	215	1	R60907	Borrelia KL11 antigen	9.22e+00
26	45	54.9	217	1	R60898	Borrelia M57 antigen v	9.22e+00
27	45	54.9	219	1	R60906	Borrelia BITS antigen	9.22e+00
28	45	54.9	233	1	R75730	B. burgdorferi strain	9.22e+00
29	45	54.9	235	1	R75728	B. burgdorferi strain	9.22e+00
30	45	54.9	303	1	W02719	G-protein coupled odor	9.22e+00
31	45	54.9	303	1	R48747	G-protein coupled odor	9.22e+00
32	45	54.9	338	1	R27875	Odorant receptor clone	9.22e+00
33	45	54.9	460	1	R62042	Dialkylglycine decarbo	9.22e+00
34	45	54.9	460	1	R36724	2,2-dialkylglycine dec	9.22e+00
35	45	54.9	640	1	W20991	H. pylori inner membra	9.22e+00
36	45	54.9	805	1	R98226	Rat neuronal protein k	9.22e+00
37	44	53.7	202	1	R62775	Borrelia 297 antigen v	1.46e+01
38	44	53.7	203	1	R62774	Borrelia ZS7 antigen v	1.46e+01
39	44	53.7	218	1	R60889	Borrelia 297 antigen v	1.46e+01
40	44	53.7	238	1	R32353	Cyclophilin C.	1.46e+01
41	44	53.7	321	1	W06515	Flavobacterium GGPP sy	1.46e+01
42	44	53.7	348	1	P93306	Fungal lipase	1.46e+01
43	44	53.7	1521	1	W18226	Transcriptional co-rep	1.46e+01
44	44	53.7	2939	1	W00384	Plasmodium falciparum	1.46e+01
45	43	52.4	214	1	R49580	Sequence of truncated	2.31e+01
46	43	52.4	220	1	R49581	Sequence of endonuclea	2.31e+01
47	43	52.4	321	1	R48758	Human cytomegalovirus	2.31e+01
48	43	52.4	321	1	W02730	Human cytomegalovirus	2.31e+01
49	43	52.4	330	1	R56480	I-19 B-lymphocyte deri	2.31e+01
50	43	52.4	343	1	R60599	Hornet phospholipase D	2.31e+01
51	43	52.4	472	1	W24535	Mannose-1-phosphate tr	2.31e+01
52	43	52.4	521	1	W13648	Geraniol/nerol 10-hydr	2.31e+01
53	43	52.4	596	1	W17767	Md-alpha-E7 malathion	2.31e+01
54	43	52.4	1300	1	R34714	Bacillus subtilis srfA	2.31e+01
55	43	52.4	1482	1	R49042	NMDA receptor channel	2.31e+01
56	43	52.4	1503	1	W10424	Saccharomyces cerevisi	2.31e+01
57	43	52.4	1503	1	R67691	S. cerevisiae scaur2R	2.31e+01
58	43	52.4	1508	1	R45944	Glutamic acid receptor	2.31e+01
59	43	52.4	1508	1	R44193	Rat NMDA receptor subu	2.31e+01
60	43	52.4	1510	1	R66040	Human N-methyl-D-aspar	2.31e+01
61	43	52.4	1510	1	R92507	Human NMDA receptor R2	2.31e+01
62	43	52.4	1510	1	R80971	Human excitatory amino	2.31e+01
63	42	51.2	92	1	W13598	Monocyte chemoattracta	3.62e+01
64	42	51.2	93	1	W13599	Monocyte chemoattracta	3.62e+01
65	42	51.2	94	1	W13597	Monocyte chemoattracta	3.62e+01
66	42	51.2	95	1	W13596	Monocyte chemoattracta	3.62e+01
67	42	51.2	102	1	R87680	Monocyte chemotactic a	3.62e+01
68	42	51.2	102	1	R87675	(28-Asp) MCP-1.	3.62e+01
69	42	51.2	102	1	R28660	MCF.	3.62e+01
70	42	51.2	110	1	R53398	Sense MCP-1.	3.62e+01
71	42	51.2	122	1	W20126	H. pylori secreted or	3.62e+01
72	42	51.2	125	1	R28663	MCF.	3.62e+01
73	42	51.2	131	1	W20940	H. pylori secreted or	3.62e+01
74	42	51.2	185	1	R12127	APP haemolysin activat	3.62e+01
75	42	51.2	211	1	W25771	Mouse mD52 C1.	3.62e+01
76	42	51.2	246	1	R80553	S-SPV-001 potential eu	3.62e+01
77	42	51.2	480	1	R97670	Human Yes-associated p	3.62e+01
78	42	51.2	496	1	R97672	Mouse Yes-associated p	3.62e+01
79	42	51.2	566	1	R22363	GroEL-1 protein.	3.62e+01
80	42	51.2	1048	1	P50097	Fusion protein of 59-V	3.62e+01
81	42	51.2	1111	1	P61048	A.nidulans phosphenol	3.62e+01
82	42	51.2	1225	1	R26188	HVTA antigen.	3.62e+01
83	42	51.2	1243	1	R49832	Beat-galactosidase/hep	3.62e+01
84	42	51.2	1280	1	W24575	Merozoite apical-end p	3.62e+01
85	42	51.2	1315	1	R76113	PS17b acaride-active t	3.62e+01
86	42	51.2	1315	1	R29027	Bacillus thuringiensis	3.62e+01
87	42	51.2	1315	1	R29517	BT toxin 17b.	3.62e+01
88	42	51.2	1315	1	R28810	BT toxin 17b.	3.62e+01
89	42	51.2	1315	1	R28890	Toxin 17b.	3.62e+01
90	42	51.2	1315	1	R20067	B.thuringiensis toxin	3.62e+01
91	42	51.2	1315	1	R58632	Bacillus thuringiensis	3.62e+01
92	42	51.2	1324	1	R85937	Protein tyrosine-kinas	3.62e+01
93	42	51.2	1360	1	R50290	lktA::lacZ fusion prot	3.62e+01
94	42	51.2	1360	1	R14481	LktA:lacZ fusion produ	3.62e+01
95	42	51.2	1411	1	R29516	BT toxin 17a.	3.62e+01
96	42	51.2	1411	1	R29026	Bacillus thuringiensis	3.62e+01

97	42	51.2	1411	1	R44201	Bacillus thuringiensis	3.62e+01
98	42	51.2	1411	1	R20066	B.thuringiensis toxin	3.62e+01
99	42	51.2	1411	1	R28803	BT toxin 17a.	3.62e+01
100	42	51.2	1411	1	R58631	Bacillus thuringiensis	3.62e+01

ALIGNMENTS

RESULT 1
ID R80619 standard; Protein; 12 AA.
AC R80619;
DT 28-FEB-1996 (first entry)
DE Immunogenic peptide of tumour rejection antigen (MAGE-1).
KW Tumour rejection antigen; MAGE-1; monoclonal antibody; MAb;
KW diagnosis; immunoassay; cancer; immunogen; antisera.
OS Homo sapiens.
PN WO9520974-A1.
PD 10-AUG-1995.
PF 05-JAN-1995; U00095.
PR 01-FEB-1994; US-190411.
PA (LUDW-) LUDWIG INST CANCER RES.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
PA (SLOK) MEMORIAL SLOAN-KETTERING CANCER CENT.
PI Boon-falleur T, Chen Y, Garin-chesa P, Old LJ, Rettig WJ;
PI Stockert E, Van der bruggen P;
DR WPI; 95-283606/37.
PT New monoclonal antibody binding specifically to MAGE-1 - useful for
PT diagnosis and monitoring of cancer, also new hybridomas, recombinant
PT MAGE-1 and immunogenic peptide(s)
PS Claim 12; Page 20; 33pp; English.
CC A monoclonal antibody directed against the tumour rejection antigen
CC (MAGE-1) can be used to detect MAGE-1 in samples by standard
CC immunoassay methods for diagnosis and monitoring of cancer etc. The
CC monoclonal antibody is designated MA454 and is produced by the
CC hybridoma deposited as ATCC HBB11540. The monoclonal antibody is
CC specific for MAGE-1, having no reactivity for MAGE-2 or MAGE-3.
CC Peptide fragments of MAGE-1 (See R80618-20) may be useful as
CC immunogens for production of the monoclonal antibody and antisera.
SQ Sequence 12 AA;

Query Match 100.0%; Score 82; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.60e-08;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 LFRAVITKKVAD 38
||||||||||||
Qy 1 LFRAVITKKVAD 12

RESULT 2
ID R70909 standard; Protein; 309 AA.
AC R70909;
DT 09-OCT-1995 (first entry)
DE Human melanoma antigen MAGE-1.
KW Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
KW HLA-restricted cytotoxic T-lymphocyte activity.
OS Homo sapiens.
PN WO9504542-A.
PD 16-FEB-1995.
PF 02-AUG-1994; U08721.
PR 06-AUG-1993; US-103623.
PA (CYTE-) CYTEL CORP.
PI Fikes JD, Livingston BD, Sette AD, Sidney JC;
DR WPI; 95-090681/12.
DR N-PSDB; Q85435.
PT Human melanoma antigen, MAGE-1, peptide(s) - useful for
PT stimulating immune response against melanoma
PS Example 1; Fig 1; 59pp; English.
CC Q85435 encodes R70909 human melanoma antigen MAGE-1, it was used
CC to produce the C-terminal MAGE-1 peptides described in R70915 to
CC R70969. These peptides are useful for defining epitopes that
CC engender a HLA-restricted cytotoxic lymphocyte activity against
CC MAGE-1 antigens. Compsns. containing these peptides can be
CC administered, as a vaccine to patients susceptable to MAGE
CC associated tumours, e.g. melanomas.
SQ Sequence 309 AA;

Query Match 100.0%; Score 82; DB 1; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.60e-08;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 123 LFRAVITKKVAD 134
||||||||||||
Qy 1 LFRAVITKKVAD 12

RESULT 3
ID R47325 standard; Protein; 10 AA.
AC R47325;
DT 31-AUG-1994 (first entry)
DE HLA-A3 MAGE 1 antigen peptide fragment 96-105.
KW Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW immune response; viral infection; cancer; prostate cancer; lymphoma;
KW hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS Synthetic.
PN WO9403205-A.
PD 17-FEB-1994.
PF 06-AUG-1993; U07421.
PR 07-AUG-1992; US-926666.
PR 05-MAR-1993; US-027746.
PA (CYTE-) CYTEL CORP.
PI Celis E, Grey HM, Kubo RT, Sette A;
DR WPI; 94-065403/08.
PT Peptide which specifically binds selected MHC allele - used to
PT induce an immune response for treatment or prevention of viral
PT infection or cancer, or for diagnosis
PS Eaxmple 8; Page 52; 150pp; English.
CC The sequences given in R47304-33 and R49201-44 are immunogenic
CC peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC These peptides may be used in the composition of the invention.
CC These peptides are capable of binding selected MHC molecules and
CC inducing an immune response. They can be used to treat and/or
CC prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC hepatitis or AIDS. They can also be used to produce antibodies for
CC use as diagnostic or therapeutic agents. The peptides can also be
CC used as diagnostic agents.
SQ Sequence 10 AA;

Query Match 79.3%; Score 65; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.04e-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 28 LFRAVITKK 36
|||||||||
Qy 1 LFRAVITKK 9

RESULT 4
ID R47324 standard; Protein; 9 AA.
AC R47324;
DT 31-AUG-1994 (first entry)
DE HLA-A3 MAGE 1 antigen peptide fragment 96-104.
KW Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW immune response; viral infection; cancer; prostate cancer; lymphoma;
KW hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS Synthetic.
PN WO9403205-A.
PD 17-FEB-1994.
PF 06-AUG-1993; U07421.
PR 07-AUG-1992; US-926666.
PR 05-MAR-1993; US-027746.
PA (CYTE-) CYTEL CORP.
PI Celis E, Grey HM, Kubo RT, Sette A;
DR WPI; 94-065403/08.
PT Peptide which specifically binds selected MHC allele - used to
PT induce an immune response for treatment or prevention of viral

PT infection or cancer, or for diagnosis
PS Eaxmple 8; Page 52; 150pp; English.
CC The sequences given in R47304-33 and R49201-44 are immunogenic
CC peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC These peptides may be used in the composition of the invention.
CC These peptides are capable of binding selected MHC molecules and
CC inducing an immune response. They can be used to treat and/or
CC prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC hepatitis or AIDS. They can also be used to produce antibodies for
CC use as diagnostic or therapeutic agents. The peptides can also be
CC used as diagnostic agents.
SQ Sequence 9 AA;

Query Match 70.7%; Score 58; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.37e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 28 LFRAVITK 35
||||||||
Qy 1 LFRAVITK 8

RESULT 5
ID R65120 standard; peptide; 9 AA.
AC R65120;
DT 09-OCT-1995 (first entry)
DE MAGE 1 immunogenic peptide 96-104.
KW MAGE 1; immunogenic peptide 96-104; cytotoxic C cells;
KW in vitro activation; cancer; AIDS; bacterial infections; malaria;
KW fungal infections; tuberculosis; hepatitis.
OS Homo sapiens.
PN WO9504817-A.
PD 16-FEB-1995.
PF 01-AUG-1994; U08672.
PR 06-AUG-1993; US-103401.
PA (CYTE-) CYTEL CORP.
PI Celis E, Kubo R, Serra H, Tsai V, Wentworth P;
DR WPI; 95-090895/12.
PT In vitro activation of cytotoxic T cells for selected killing of
PT target cells - for treating e.g. cancer, AIDS, hepatitis etc.by
PT incubating them with antigen presenting cells loaded with
PT appropriate immunogenic peptide
PS Example 3; Page 35; 53pp; English.
CC R65109-R65145 are immunogenic peptides, they are used in a new
CC method for the in vitro activation of cytotoxic T cells (CTC).
CC This is achieved by incubating the CTCs with antigen presenting
CC cells loaded with an appropriate immunogenic peptide (e.g. one
CC of the above peptides). By selecting the peptides used the
CC following diseases and infections can be treated; cancer, AIDS,
CC hepatitis, other viral and bacterial infections, malaria and
CC tuberculosis.
SQ Sequence 9 AA;

Query Match 70.7%; Score 58; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.37e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 28 LFRAVITK 35
||||||||
Qy 1 LFRAVITK 8

RESULT 6
ID R49228 standard; Protein; 9 AA.
AC R49228;
DT 31-AUG-1994 (first entry)
DE HLA-A11 MAGE 1 antigen peptide fragment 1072.13.
KW Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW immune response; viral infection; cancer; prostate cancer; lymphoma;
KW hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS Synthetic.
PN WO9403205-A.
PD 17-FEB-1994.
PF 06-AUG-1993; U07421.
PR 07-AUG-1992; US-926666.
PR 05-MAR-1993; US-027746.
PA (CYTE-) CYTEL CORP.
PI Celis E, Grey HM, Kubo RT, Sette A;
DR WPI; 94-065403/08.
PT Peptide which specifically binds selected MHC allele - used to
PT induce an immune response for treatment or prevention of viral
PT infection or cancer, or for diagnosis
PS Eaxmple 16; Page 116; 150pp; English.
CC The sequences given in R47304-33 and R49201-44 are immunogenic
CC peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC These peptides may be used in the composition of the invention.
CC These peptides are capable of binding selected MHC molecules and
CC inducing an immune response. They can be used to treat and/or
CC prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC hepatitis or AIDS. They can also be used to produce antibodies for
CC use as diagnostic or therapeutic agents. The peptides can also be
CC used as diagnostic agents.
SQ Sequence 9 AA;

Query Match 70.7%; Score 58; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.37e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 28 LFRAVITK 35
||||||||
Qy 1 LFRAVITK 8

RESULT 7
ID R65125 standard; peptide; 10 AA.
AC R65125;
DT 09-OCT-1995 (first entry)
DE MAGE 1 immunogenic peptide 95-104.
KW MAGE 1; immunogenic peptide 95-104; cytotoxic C cells;
KW in vitro activation; cancer; AIDS; bacterial infections; malaria;
KW fungal infections; tuberculosis; hepatitis.
OS Homo sapiens.
PN WO9504817-A.
PD 16-FEB-1995.
PF 01-AUG-1994; U08672.
PR 06-AUG-1993; US-103401.
PA (CYTE-) CYTEL CORP.
PI Celis E, Kubo R, Serra H, Tsai V, Wentworth P;
DR WPI; 95-090895/12.
PT In vitro activation of cytotoxic T cells for selected killing of
PT target cells - for treating e.g. cancer, AIDS, hepatitis etc.by
PT incubating them with antigen presenting cells loaded with
PT appropriate immunogenic peptide
PS Example 3; Page 35; 53pp; English.
CC R65109-R65145 are immunogenic peptides, they are used in a new
CC method for the in vitro activation of cytotoxic T cells (CTC).
CC This is achieved by incubating the CTCs with antigen presenting
CC cells loaded with an appropriate immunogenic peptide (e.g. one
CC of the above peptides). By selecting the peptides used the
CC following diseases and infections can be treated; cancer, AIDS,
CC hepatitis, other viral and bacterial infections, malaria and
CC tuberculosis.
SQ Sequence 10 AA;

Query Match 70.7%; Score 58; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.37e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 29 LFRAVITK 36
||||||||
Qy 1 LFRAVITK 8

RESULT 8
ID R49230 standard; Protein; 10 AA.
AC R49230;

```
DT   31-AUG-1994  (first entry)
DE   HLA-A11 MAGE 1 antigen peptide fragment 1072.15.
KW   Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW   immune response; viral infection; cancer; prostate cancer; lymphoma;
KW   hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS   Synthetic.
PN   WO9403205-A.
PD   17-FEB-1994.
PF   06-AUG-1993; U07421.
PR   07-AUG-1992; US-926666.
PR   05-MAR-1993; US-027746.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Grey HM,  Kubo RT,  Sette A;
DR   WPI; 94-065403/08.
PT   Peptide which specifically binds selected MHC allele - used to
PT   induce an immune response for treatment or prevention of viral
PT   infection or cancer, or for diagnosis
PS   Eaxmple 16; Page 116; 150pp; English.
CC   The sequences given in R47304-33 and R49201-44 are immunogenic
CC   peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC   These peptides may be used in the composition of the invention.
CC   These peptides are capable of binding selected MHC molecules and
CC   inducing an immune response.  They can be used to treat and/or
CC   prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC   hepatitis or AIDS.  They can also be used to produce antibodies for
CC   use as diagnostic or therapeutic agents.  The peptides can also be
CC   used as diagnostic agents.
SQ   Sequence   10 AA;

  Query Match           70.7%;  Score 58;  DB 1;  Length 36;
  Best Local Similarity 100.0%;  Pred. No. 1.37e-02;
  Matches    8;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     29 LFRAVITK 36
          ||||||||
Qy      1 LFRAVITK 8


RESULT   9
ID   W11578 standard; Protein; 578 AA.
AC   W11578;
DT   25-MAR-1997  (first entry)
DE   Bacillus licheniformis maltogenic amylase.
KW   BLMA; amyolytic enzyme; maltogenic amylase; alpha-1,6-linkage;
KW   branched oligosaccharide; pullulan; cyclodextrin; hydrolysis;
KW   sugar transferase activity.
OS   Bacillus licheniformis (ATCC 27811).
FH   Key             Location/Qualifiers
FT   Region          238. .244
FT   /label= conserved_motif
FT   /note= "putative calcium binding domain"
FT   Region          318. .326
FT   /label= conserved_motif
FT   /note= "sequence conserved among amylases"
FT   Region          349. .352
FT   /label= conserved_motif
FT   /note= "sequence conserved among amylases"
FT   Region          411. .416
FT   /label= conserved_motif
FT   /note= "sequence conserved among amylases"
PN   US5583039-A.
PD   10-DEC-1996.
PF   07-JUN-1990; 534679.
PR   07-JUN-1990; US-534679.
PR   12-NOV-1993; US-152271.
PA   (DOOS-) DOOSAN TECH CENT.
PA   (SUNH-) SUNHILL GLUCOSE CO LTD.
PA   (SUNK-) SUNKYONG IND LTD.
PI   Cha JH,  Choi YD,  Jang SY,  Kim IC,  Kim JR,  Kim KH;
PI   Park KH,  Seo BC;
DR   WPI; 97-042316/04.
DR   N-PSDB; T58548.
PT   Recombinant E. coli strain - contg. gene coding for Bacillus
PT   licheniformis maltogenic amylase
PS   Claim 2; Fig 2; 21pp; English.
CC   A genomic library was constructed from EcoRI- and BamHI-partially
CC   digested chromosomal DNA of Bacillus licheniformis. The DNA
CC   fragments were cloned into plasmid pBR322 and transformed into
CC   E.coli HB101. Ampicillin-resistant colonies were screened for
CC   amyolytic activity on starch agar. One positive colony showed starch
CC   hydrolysing phenotype after the cell membrane was disrupted
CC   with D-cycloserine. The DNA insert from positive transformants was
CC   3.5 kb long. A subfragment containing an open reading frame of 1740
CC   nucleotides was sequenced. The gene product has the present deduced
CC   amino acid sequence and is a B.licheniformis maltogenic amylase
CC   (BLMA) which hydrolyses starch, pullulan and cyclodextrin at an optimum
CC   temperature of 50 deg.C and pH 7. The enzyme also has sugar
CC   transferase activity in the presence of glucose. A previously
CC   reported BLMA enzyme was thermostable and only hydrolysed starch.
CC   The new BLMA can be used for synthesising branched oligosaccharides
CC   with alpha-1,6-linkages.
SQ   Sequence   578 AA;

  Query Match           58.5%;  Score 48;  DB 1;  Length 604;
  Best Local Similarity 50.0%;  Pred. No. 2.22e+00;
  Matches    4;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    247 LFRTVVSR 254
          |||:|:::
Qy      1 LFRAVITK 8


RESULT   10
ID   R67916 standard; Protein; 309 AA.
AC   R67916;
DT   14-SEP-1995  (first entry)
DE   (1-3)-beta-D-glucan sensitive factor.
KW   (1-3)-beta-D-glucan sensitive factor; antifungal agent;
KW   mycosis diagnosis.
OS   Limulus sp.
FH   Key             Location/Qualifiers
FT   Peptide         1. .31
FT   /label= sig_peptide
PN   WO9501432-A.
PD   12-JAN-1995.
PF   29-JUN-1994; J01057.
PR   29-JUN-1993; JP-184403.
PA   (SEGK ) SEIKAGAKU KOGYO CO LTD.
PI   Iwanaga S,  Muta T,  Oda T,  Seki N;
DR   WPI; 95-060996/08.
DR   N-PSDB; Q81335.
PT   DNA encoding a polypeptide comprising a tetrapeptide motif at
PT   least once - which may be used as an antibacterial and
PT   antifungal.
PS   Claim 11; Pages 32-38; 51pp; Japanese.
CC   Q81335 encodes R67916 a (1-3)-beta-D-glucan sensitive factor, it
CC   has a high affinity for the (1-3)-beta-D-glucan found in fungal
CC   cell walls. The protein is therefore useful for clinically
CC   diagnosing mycosis, and as an antifungal agent for the removal
CC   of fungi.
SQ   Sequence   309 AA;

  Query Match           57.3%;  Score 47;  DB 1;  Length 335;
  Best Local Similarity 50.0%;  Pred. No. 3.59e+00;
  Matches    5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db     66 FRPVITRIIG 75
          ||:|||: ::
Qy      2 FRAVITKKVA 11


RESULT   11
ID   R98227 standard; Protein; 722 AA.
AC   R98227;
DT   23-SEP-1996  (first entry)
DE   Rat neuronal protein kinase MARK-2.
```

```
KW   Neuronal protein kinase; NPK; microtubule associated protein;
KW   MAP; tau protein; phosphorylation; NPK inhibitor; Alzheimer disease;
KW   cancer; therapy; diagnosis.
OS   Rattus norvegicus.
PN   WO9613592-A2.
PD   09-MAY-1996.
PF   30-OCT-1995; E04258.
PR   28-OCT-1994; EP-117122.
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PI   Biernat J,  Drewes G,  Mandelkow E;
DR   WPI; 96-251461/25.
PT   DNA encoding neuronal protein kinase (NPK) - useful for identifying
PT   NPK inhibitors for treatment of Alzheimer's disease and cancer.
PS   Claim 1; Page 44-45; 77pp; English.
CC   A novel rat neuronal protein kinase (R98227), designated NPK MARK-2,
CC   is capable of phosphorylating a KXGS sequence motif in tau protein
CC   and microtubule associated proteins MAP4, MAP2 and MAP2c (see also
CC   R98229-39 and W00850-54), causing their dissociation from microtubules.
CC   Phosphorylation of human tau Ser-262 is indicative of the onset of
CC   Alzheimer's disease.  MARK-1 is the product of a cDNA clone obtd.
CC   from a rat brain cDNA library by screening with probes derived
CC   from pig brain peptide sequences (see also R98240-50).  Another NPK,
CC   MARK-1 (R98226), was similarly isolated.  Inhibitors (e.g. antibodies)
CC   of NPKs are used to treat Alzheimer's disease and cancer.  NPKs are
CC   themselves used for in vitro diagnosis and/or monitoring of
CC   Alzheimer's disease and cancer.
SQ   Sequence   722 AA;

  Query Match           57.3%;  Score 47;  DB 1;  Length 748;
  Best Local Similarity  60.0%;  Pred. No. 3.59e+00;
  Matches     6;  Conservative    1;  Mismatches   3;  Indels   0;  Gaps   0;

Db    123 LFREVRIMKV 132
          ||| |  :||
Qy      1 LFRAVITKKV 10


RESULT   12
ID   P91891 standard; protein; 140 AA.
AC   P91891;
DT   29-APR-1990  (first entry)
DE   Part of the sequence of the Brazil nut 2S-albumin as encoded in the
DE   pBN2S1 plasmid
KW   2S-albumin; brazil nut; pBN2S1; storage protein gene;
KW   heterologous polypeptide.
OS   Brazil nut.
FH   Key             Location/Qualifiers
FT   peptide         1. .30
FT   /note="signal peptide"
FT   protein         31. .38
FT   /note="mature small subunit"
FT   region          39. .43
FT   /note="processing site"
FT   protein         44. .136
FT   /note="mature large subunit"
PN   WO8903887-A.
PD   05-MAY-1989.
PF   20-OCT-1988; E00944.
PR   20-OCT-1987; EP-402348.
PA   (PLAN-) Plant Genetic Syst.
PI   Vandekerckhove JS, Krebbers E, Botterman J, Leemans J;
DR   WPI; 89-150783/20.
DR   N-PSDB; N91699.
PT   Recombinant DNA expression in plants
PT   - using modified storage protein genes for expressing
PT   heterologous polypeptide(s) in the seeds
PS   Figure 4; 121pp; English.
CC   The entire 2S-albumin storage protein precursor including
CC   signal peptide. It is to be inserted into plants under the control of
CC   a seed-specific promoter and expressed at high levels only or mostly
CC   in the seed forming stage and produced mostly in the seeds.
SQ   Sequence   140 AA;

  Query Match           56.1%;  Score 46;  DB 1;  Length 166;
  Best Local Similarity  45.5%;  Pred. No. 5.77e+00;
  Matches     5;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db     43 FRATVTTTVVE 53
          ||| :|  | :
Qy      2 FRAVITKKVAD 12


RESULT   13
ID   P81996 standard; protein; 330 AA.
AC   P81996;
DT   17-DEC-1990  (first entry)
DE   Sequence encoded by nodulation regulatory gene 2 (nodD-2) of
DE   Bradyrhizobium japonicum strain USDA 123
KW   Rhizobium; symbiosis; legume.
OS   Bradyrhizobium japonicum USDA 123.
PN   WO8707910-A.
PD   30-DEC-1987.
PF   17-JUN-1987; U01421.
PR   7-JUN-1986; US-875297.
PR   11-JUN-1987; US-061848.
PA   (LUBR) Lubrizol Genetics I.
PI   Appelbaum ER, Hennecke H, Lamb JW, Gottfert M;
DR   WPI; 88-014399/02.
DR   N-PSDB; N82007.
PT   DNA contg. nodulation regulatory genes (nod D) -
PT   from Bradyrhizobium japonicum strains, useful for selective
PT   expression of structural genes
PS   Disclosure;  4-4; 88pp; English.
CC   DNA sequence of the nodD-1 gene coding regionof B.japonicum USDA 110 is
CC   almost identical to that of USDA 123 (N82006) with single base change in
CC   codon 139 which is GAC in USDA 110. The deduced protein sequences of the
CC   two nodD-1 differ by one AA at position 139, which is Asp in USDA 110.
CC   nodD-2 may be used to enhance competitiveness of strains for nodulation
CC   and in the selective manipulation of nodulation host range of strains.
SQ   Sequence   330 AA;

  Query Match           56.1%;  Score 46;  DB 1;  Length 356;
  Best Local Similarity  50.0%;  Pred. No. 5.77e+00;
  Matches     4;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    137 LFRNVVAR 144
          ||| |:::
Qy      1 LFRAVITK 8


RESULT   14
ID   W07702 standard; Protein; 543 AA.
AC   W07702;
DT   06-APR-1997  (first entry)
DE   Mouse ETS2 repressor factor (ERF).
KW   ETS2 repressor factor; ERF; transcriptional repressor;
KW   tumour suppressor; tumour; cancer; oncoprotein; therapy.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   Domain          21. .98
FT   /label= DNA_binding_domain
FT   /note= "ets-like DNA binding domain"
FT   Domain          466. .525
FT   /label= Active_repressor_domain
PN   WO9639517-A1.
PD   12-DEC-1996.
PF   04-JUN-1996; U10177.
PR   05-JUN-1995; US-469412.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Athanasiou MA,  Beal GJ,  Blair DG,  Fisher RJ,  Mavrothalassitis GJ;
PI   Sgouras D N;
DR   WPI; 97-043139/04.
DR   N-PSDB; T47200.
PT   New DNA encoding ETS2 repressor factor - useful for reducing
PT   tumourigenicity, esp. oncogene associated tumour cells
PS   Disclosure; Page 70-72; 101pp; English.
```

```
CC   Murine ETS2 repressor factor (ERF) (W07702) is a member of the ETS
CC   family and acts as a transcriptional repressor in mammalian cells.
CC   Its amino acid sequence was deduced from the murine ERF gene
CC   (T47198).  Human ERF (see also W07700) has also been identified.
CC   ERF has tumour suppressor activity.  Chimeric molecules comprising
CC   the ERF repressor domain in combination with a heterologous
CC   transcription factor having a binding domain can be used to reduce
CC   tumourigenicity associated with inappropriate expression of
CC   transcription factors.
SQ   Sequence   543 AA;

  Query Match           56.1%;  Score 46;  DB 1;  Length 569;
  Best Local Similarity  36.4%;  Pred. No. 5.77e+00;
  Matches    4;  Conservative    6;  Mismatches   1;  Indels   0;  Gaps   0;

Db    190 LFSAVVARRLG 200
          || ||::::::
Qy      1 LFRAVITKKVA 11


RESULT   15
ID   W07700 standard; Protein; 548 AA.
AC   W07700;
DT   06-APR-1997  (first entry)
DE   Human ETS2 repressor factor (ERF).
KW   ETS2 repressor factor; ERF; transcriptional repressor;
KW   tumour suppressor; tumour; cancer; oncoprotein; therapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          29. .106
FT   /label= DNA_binding_domain
FT   /note= "ets-like DNA binding domain"
FT   Domain          472. .530
FT   /label= Active_repressor_domain
FT   /note= "(Claim 20)"
PN   WO9639517-A1.
PD   12-DEC-1996.
PF   04-JUN-1996; U10177.
PR   05-JUN-1995; US-469412.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Athanasiou MA,  Beal GJ,  Blair DG,  Fisher RJ,  Mavrothalassitis GJ;
PI   Sgouras D N;
DR   WPI; 97-043139/04.
DR   N-PSDB; T47198.
PT   New DNA encoding ETS2 repressor factor - useful for reducing
PT   tumourigenicity, esp. oncogene associated tumour cells
PS   Claim 1; Page 59-61; 101pp; English.
CC   Novel human ETS2 repressor factor (ERF) (W07700) is the first member
CC   of the ETS family to be identified as a transcriptional repressor in
CC   mammalian cells.  Its amino acid sequence was deduced from a cDNA
CC   clone (T47198) derived from K562 cells.  ERF and alternatively
CC   spliced ERF (see also W07701) show no homology to other known
CC   proteins.  The ERF repressor domain in combination with a
CC   heterologous transcription factor having a binding domain can
CC   be used as novel transcriptional repressors to reduce
CC   tumourigenicity associated with inappropriate expression of the
CC   GAL4, NFkappaB (HIV), MYC (Burkitt lymphoma), Fli-1 (Ewing's
CC   sarcoma) and EST1 transcription factors.
SQ   Sequence   548 AA;

  Query Match           56.1%;  Score 46;  DB 1;  Length 574;
  Best Local Similarity  36.4%;  Pred. No. 5.77e+00;
  Matches    4;  Conservative    6;  Mismatches   1;  Indels   0;  Gaps   0;

Db    198 LFSAVVARRLG 208
          || ||::::::
Qy      1 LFRAVITKKVA 11


RESULT   16
ID   R90617 standard; Protein; 728 AA.
AC   R90617;
DT   29-JUN-1996  (first entry)
DE   Sulfolobus solfataricus transferase for alpha, alpha-trehalose prodn.
KW   transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;
KW   malto-oligosaccharide; hydrolysis.
OS   Sulfolobus solfataricus.
PN   WO9534642-A.
PD   21-DEC-1995.
PF   14-JUN-1995; J01189.
PR   15-JUN-1994; JP-133354.
PR   18-AUG-1994; JP-194223.
PR   31-OCT-1994; JP-290394.
PR   21-NOV-1994; JP-311185.
PR   21-NOV-1994; JP-286917.
PR   21-APR-1995; JP-120673.
PA   (KIRI ) KIRIN BEER KK.
DR   WPI; 96-049671/05.
DR   N-PSDB; T12323.
PT   Sulfolobus spp. derived transferase and amylase - for production of
PT   alpha, alpha-trehalose from malto-oligosaccharide(s)
PS   Claim 74; Page 213-219; 357pp; Japanese.
CC   The transferase is derived from Sulfolobus solfataricus.  The transferase
CC   acts on a saccharide having at least three sugar units, in which at least
CC   three glucose units at the reducing end are alpha-1,4 linked, to
CC   transform the alpha-1,4 linkages to alpha-1, alpha-1 linkages. The
CC   transferase has a mol. wt. of 74 to 76 kDa.  It is characterised by
CC   working at pH 4.5-6.0 and at 60-80 deg.C.  It has an isoelectric point
CC   of 5.3-6.3 and retains at least 90 percent activity after 6 hrs. at 80
CC   deg.C. It is completely inhibited by 5 mM copper sulphate. Use of the
CC   transferase and an amylase in succession on suitable substrates such
CC   a malto-oligosaccharides, is useful for the production of
CC   alpha, alpha-trehalose.
SQ   Sequence   728 AA;

  Query Match           56.1%;  Score 46;  DB 1;  Length 754;
  Best Local Similarity  33.3%;  Pred. No. 5.77e+00;
  Matches    4;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db    702 LFSPIVTREVKE 713
          || :::|: | :
Qy      1 LFRAVITKKVAD 12


RESULT   17
ID   R60900 standard; Protein; 159 AA.
AC   R60900;
DT   25-MAY-1995  (first entry)
DE   Borrelia VSDA antigen vaccine.
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris; ss.
OS   Borrelia burgdorferi VSDA.
PN   WO9425596-A.
PD   10-NOV-1994.
PF   29-APR-1994; E01365.
PR   29-APR-1993; US-053863.
PA   (IMMO ) IMMUNO AG.
PI   Crowe B,  Dorner F,  Livey I;
DR   WPI; 94-358273/44.
DR   N-PSDB; Q73873.
PT   Immunogenic composition comprising OspC antigens - for the
PT   treatment of Lyme borreliosis in different, specific geographical
PT   areas.
PS   Disclosure; Fig. 9a; 115pp; English.
CC   A vaccine for Lyme disease includes selected OspC antigen
CC   formulations based on defined OspC families resolved by serovar
CC   typing and RFLP typing.  Partial sequences of OspC genes selected
CC   from different RFLP types are given in Q73883-905 (encoded peptides,
CC   comprising the first 92% of mature OspC, are given in R62771-93).
CC   Complete sequences of these novel ospC genes, including the 3' end,
CC   plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC   are given in Q73857-82, and encoded proteins in R60884-909.  The
CC   DNA sequences may be expressed in e.g. Pichia pastoris for
CC   recombinant antigen production.
SQ   Sequence   159 AA;
```

```
  Query Match           54.9%;  Score 45;  DB 1;  Length 185;
  Best Local Similarity  55.6%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     49 TVISKKITD 57
          :||:||::|
Qy      4 AVITKKVAD 12


RESULT   18
ID   R62793 standard; Protein; 173 AA.
AC   R62793;
DT   25-MAY-1995  (first entry)
DE   Borrelia KL11 antigen vaccine.
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris.
OS   Borrelia burgdorferi KL11.
PN   WO9425596-A.
PD   10-NOV-1994.
PF   29-APR-1994; E01365.
PR   29-APR-1993; US-053863.
PA   (IMMO ) IMMUNO AG.
PI   Crowe B,  Dorner F,  Livey I;
DR   WPI; 94-358273/44.
DR   N-PSDB; Q73905.
PT   Immunogenic composition comprising OspC antigens - for the
PT   treatment of Lyme borreliosis in different, specific geographical
PT   areas.
PS   Disclosure; Fig. 9; 115pp; English.
CC   A vaccine for Lyme disease includes selected OspC antigen
CC   formulations based on defined OspC families resolved by serovar
CC   typing and RFLP typing.  Partial sequences of OspC genes selected
CC   from different RFLP types are given in Q73883-905 (encoded peptides,
CC   comprising the first 92% of mature OspC, are given in R62771-93).
CC   Complete sequences of these novel ospC genes, including the 3' end,
CC   plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC   are given in Q73857-82, and encoded proteins in R60884-909.  The
CC   DNA sequences may be expressed in e.g. Pichia pastoris for
CC   recombinant antigen production.
SQ   Sequence   173 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 199;
  Best Local Similarity  55.6%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     49 TVISKKITD 57
          :||:||::|
Qy      4 AVITKKVAD 12


RESULT   19
ID   R60908 standard; Protein; 173 AA.
AC   R60908;
DT   25-MAY-1995  (first entry)
DE   Borrelia PBI antigen vaccine.
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris.
OS   Borrelia burgdorferi PBI.
PN   WO9425596-A.
PD   10-NOV-1994.
PF   29-APR-1994; E01365.
PR   29-APR-1993; US-053863.
PA   (IMMO ) IMMUNO AG.
PI   Crowe B,  Dorner F,  Livey I;
DR   WPI; 94-358273/44.
DR   N-PSDB; Q73881.
PT   Immunogenic composition comprising OspC antigens - for the
PT   treatment of Lyme borreliosis in different, specific geographical
PT   areas.
PS   Disclosure; Fig. 9a; 115pp; English.
CC   A vaccine for Lyme disease includes selected OspC antigen
CC   formulations based on defined OspC families resolved by serovar
CC   typing and RFLP typing.  Partial sequences of OspC genes selected
CC   from different RFLP types are given in Q73883-905 (encoded peptides,
CC   comprising the first 92% of mature OspC, are given in R62771-93).
CC   Complete sequences of these novel ospC genes, including the 3' end,
CC   plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC   are given in Q73857-82, and encoded proteins in R60884-909.  The
CC   DNA sequences may be expressed in e.g. Pichia pastoris for
CC   recombinant antigen production.
SQ   Sequence   173 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 199;
  Best Local Similarity  55.6%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     49 TVISKKITD 57
          :||:||::|
Qy      4 AVITKKVAD 12


RESULT   20
ID   R62786 standard; Protein; 174 AA.
AC   R62786;
DT   25-MAY-1995  (first entry)
DE   Borrelia VSDA antigen vaccine.
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris.
OS   Borrelia burgdorferi VSDA.
PN   WO9425596-A.
PD   10-NOV-1994.
PF   29-APR-1994; E01365.
PR   29-APR-1993; US-053863.
PA   (IMMO ) IMMUNO AG.
PI   Crowe B,  Dorner F,  Livey I;
DR   WPI; 94-358273/44.
DR   N-PSDB; Q73898.
PT   Immunogenic composition comprising OspC antigens - for the
PT   treatment of Lyme borreliosis in different, specific geographical
PT   areas.
PS   Disclosure; Fig. 9; 115pp; English.
CC   A vaccine for Lyme disease includes selected OspC antigen
CC   formulations based on defined OspC families resolved by serovar
CC   typing and RFLP typing.  Partial sequences of OspC genes selected
CC   from different RFLP types are given in Q73883-905 (encoded peptides,
CC   comprising the first 92% of mature OspC, are given in R62771-93).
CC   Complete sequences of these novel ospC genes, including the 3' end,
CC   plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC   are given in Q73857-82, and encoded proteins in R60884-909.  The
CC   DNA sequences may be expressed in e.g. Pichia pastoris for
CC   recombinant antigen production.
SQ   Sequence   174 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 200;
  Best Local Similarity  55.6%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     49 TVISKKITD 57
          :||:||::|
Qy      4 AVITKKVAD 12


RESULT   21
ID   R62784 standard; Protein; 175 AA.
AC   R62784;
DT   25-MAY-1995  (first entry)
DE   Borrelia M57 antigen vaccine.
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris.
OS   Borrelia burgdorferi M57.
PN   WO9425596-A.
```

PD 10-NOV-1994.
PF 29-APR-1994; E01365.
PR 29-APR-1993; US-053863.
PA (IMMO) IMMUNO AG.
PI Crowe B, Dorner F, Livey I;
DR WPI; 94-358273/44.
DR N-PSDB; Q73896.
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
PS Disclosure; Fig. 9; 115pp; English.
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing. Partial sequences of OspC genes selected
CC from different RFLP types are given in Q73883-905 (encoded peptides,
CC comprising the first 92% of mature OspC, are given in R62771-93).
CC Complete sequences of these novel ospC genes, including the 3' end,
CC plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC are given in Q73857-82, and encoded proteins in R60884-909. The
CC DNA sequences may be expressed in e.g. Pichia pastoris for
CC recombinant antigen production.
SQ Sequence 175 AA;

Query Match 54.9%; Score 45; DB 1; Length 201;
Best Local Similarity 55.6%; Pred. No. 9.22e+00;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 49 TVISKKITD 57
:||:||::|
Qy 4 AVITKKVAD 12

RESULT 22
ID R60904 standard; Protein; 177 AA.
AC R60904;
DT 25-MAY-1995 (first entry)
DE Borrelia IP90 antigen vaccine.
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP; Pichia pastoris.
OS Borrelia burgdorferi IP90.
PN WO9425596-A.
PD 10-NOV-1994.
PF 29-APR-1994; E01365.
PR 29-APR-1993; US-053863.
PA (IMMO) IMMUNO AG.
PI Crowe B, Dorner F, Livey I;
DR WPI; 94-358273/44.
DR N-PSDB; Q60904.
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
PS Disclosure; Fig. 9a; 115pp; English.
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing. Partial sequences of OspC genes selected
CC from different RFLP types are given in Q73883-905 (encoded peptides,
CC comprising the first 92% of mature OspC, are given in R62771-93).
CC Complete sequences of these novel ospC genes, including the 3' end,
CC plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC are given in Q73857-82, and encoded proteins in R60884-909. The
CC DNA sequences may be expressed in e.g. Pichia pastoris for
CC recombinant antigen production.
SQ Sequence 177 AA;

Query Match 54.9%; Score 45; DB 1; Length 203;
Best Local Similarity 55.6%; Pred. No. 9.22e+00;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 49 TVISKKITD 57
:||:||::|
Qy 4 AVITKKVAD 12

RESULT 23
ID R62792 standard; Protein; 177 AA.
AC R62792;
DT 25-MAY-1995 (first entry)
DE Borrelia BITS antigen vaccine.
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP; Pichia pastoris.
OS Borrelia burgdorferi BITS.
PN WO9425596-A.
PD 10-NOV-1994.
PF 29-APR-1994; E01365.
PR 29-APR-1993; US-053863.
PA (IMMO) IMMUNO AG.
PI Crowe B, Dorner F, Livey I;
DR WPI; 94-358273/44.
DR N-PSDB; Q73904.
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
PS Disclosure; Fig. 9; 115pp; English.
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing. Partial sequences of OspC genes selected
CC from different RFLP types are given in Q73883-905 (encoded peptides,
CC comprising the first 92% of mature OspC, are given in R62771-93).
CC Complete sequences of these novel ospC genes, including the 3' end,
CC plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC are given in Q73857-82, and encoded proteins in R60884-909. The
CC DNA sequences may be expressed in e.g. Pichia pastoris for
CC recombinant antigen production.
SQ Sequence 177 AA;

Query Match 54.9%; Score 45; DB 1; Length 203;
Best Local Similarity 55.6%; Pred. No. 9.22e+00;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 49 TVISKKITD 57
:||:||::|
Qy 4 AVITKKVAD 12

RESULT 24
ID R62790 standard; Protein; 177 AA.
AC R62790;
DT 25-MAY-1995 (first entry)
DE Borrelia IP90 antigen vaccine.
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP; Pichia pastoris.
OS Borrelia burgdorferi IP90.
PN WO9425596-A.
PD 10-NOV-1994.
PF 29-APR-1994; E01365.
PR 29-APR-1993; US-053863.
PA (IMMO) IMMUNO AG.
PI Crowe B, Dorner F, Livey I;
DR WPI; 94-358273/44.
DR N-PSDB; Q73902.
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
PS Disclosure; Fig. 9; 115pp; English.
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing. Partial sequences of OspC genes selected
CC from different RFLP types are given in Q73883-905 (encoded peptides,
CC comprising the first 92% of mature OspC, are given in R62771-93).
CC Complete sequences of these novel ospC genes, including the 3' end,
CC plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC are given in Q73857-82, and encoded proteins in R60884-909. The
CC DNA sequences may be expressed in e.g. Pichia pastoris for

```
CC   recombinant antigen production.
SQ   Sequence   177 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 203;
  Best Local Similarity  55.6%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db      49 TVISKKITD 57
           :||:||::|
Qy       4 AVITKKVAD 12


RESULT   25
ID   R60907 standard; Protein; 189 AA.
AC   R60907;
DT   25-MAY-1995  (first entry)
DE   Borrelia KL11 antigen vaccine.
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris.
OS   Borrelia burgdorferi KL11.
PN   WO9425596-A.
PD   10-NOV-1994.
PF   29-APR-1994; E01365.
PR   29-APR-1993; US-053863.
PA   (IMMO ) IMMUNO AG.
PI   Crowe B,  Dorner F,  Livey I;
DR   WPI; 94-358273/44.
DR   N-PSDB; Q73880.
PT   Immunogenic composition comprising OspC antigens - for the
PT   treatment of Lyme borreliosis in different, specific geographical
PT   areas.
PS   Disclosure; Fig. 9a; 115pp; English.
CC   A vaccine for Lyme disease includes selected OspC antigen
CC   formulations based on defined OspC families resolved by serovar
CC   typing and RFLP typing.  Partial sequences of OspC genes selected
CC   from different RFLP types are given in Q73883-905 (encoded peptides,
CC   comprising the first 92% of mature OspC, are given in R62771-93).
CC   Complete sequences of these novel ospC genes, including the 3' end,
CC   plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC   are given in Q73857-82, and encoded proteins in R60884-909.  The
CC   DNA sequences may be expressed in e.g. Pichia pastoris for
CC   recombinant antigen production.
SQ   Sequence   189 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 215;
  Best Local Similarity  55.6%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db      49 TVISKKITD 57
           :||:||::|
Qy       4 AVITKKVAD 12


RESULT   26
ID   R60898 standard; Protein; 191 AA.
AC   R60898;
DT   25-MAY-1995  (first entry)
DE   Borrelia M57 antigen vaccine.
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris.
OS   Borrelia burgdorferi M57.
PN   WO9425596-A.
PD   10-NOV-1994.
PF   29-APR-1994; E01365.
PR   29-APR-1993; US-053863.
PA   (IMMO ) IMMUNO AG.
PI   Crowe B,  Dorner F,  Livey I;
DR   WPI; 94-358273/44.
DR   N-PSDB; Q73871.
PT   Immunogenic composition comprising OspC antigens - for the
PT   treatment of Lyme borreliosis in different, specific geographical
PT   areas.
PS   Disclosure; Fig. 9a; 115pp; English.
CC   A vaccine for Lyme disease includes selected OspC antigen
CC   formulations based on defined OspC families resolved by serovar
CC   typing and RFLP typing.  Partial sequences of OspC genes selected
CC   from different RFLP types are given in Q73883-905 (encoded peptides,
CC   comprising the first 92% of mature OspC, are given in R62771-93).
CC   Complete sequences of these novel ospC genes, including the 3' end,
CC   plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC   are given in Q73857-82, and encoded proteins in R60884-909.  The
CC   DNA sequences may be expressed in e.g. Pichia pastoris for
CC   recombinant antigen production.
SQ   Sequence   191 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 217;
  Best Local Similarity  55.6%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db      49 TVISKKITD 57
           :||:||::|
Qy       4 AVITKKVAD 12


RESULT   27
ID   R60906 standard; Protein; 193 AA.
AC   R60906;
DT   25-MAY-1995  (first entry)
DE   Borrelia BITS antigen vaccine.
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris.
OS   Borrelia burgdorferi BITS.
PN   WO9425596-A.
PD   10-NOV-1994.
PF   29-APR-1994; E01365.
PR   29-APR-1993; US-053863.
PA   (IMMO ) IMMUNO AG.
PI   Crowe B,  Dorner F,  Livey I;
DR   WPI; 94-358273/44.
DR   N-PSDB; Q73879.
PT   Immunogenic composition comprising OspC antigens - for the
PT   treatment of Lyme borreliosis in different, specific geographical
PT   areas.
PS   Disclosure; Fig. 9a; 115pp; English.
CC   A vaccine for Lyme disease includes selected OspC antigen
CC   formulations based on defined OspC families resolved by serovar
CC   typing and RFLP typing.  Partial sequences of OspC genes selected
CC   from different RFLP types are given in Q73883-905 (encoded peptides,
CC   comprising the first 92% of mature OspC, are given in R62771-93).
CC   Complete sequences of these novel ospC genes, including the 3' end,
CC   plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC   are given in Q73857-82, and encoded proteins in R60884-909.  The
CC   DNA sequences may be expressed in e.g. Pichia pastoris for
CC   recombinant antigen production.
SQ   Sequence   193 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 219;
  Best Local Similarity  55.6%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db      49 TVISKKITD 57
           :||:||::|
Qy       4 AVITKKVAD 12


RESULT   28
ID   R75730 standard; Protein; 207 AA.
AC   R75730;
DT   31-JUL-1996  (first entry)
DE   B. burgdorferi strain pTrob outer surface protein C (OspC-pTrob).
KW   Strain pTrob; outer surface protein; OspC; antigenic domain;
KW   chimaeric protein; treatment; diagnosis; infection; vaccine;
KW   Lyme borreliosis; immunodiagnostic assay; antibody;
```

KW T-cell reactivity; chimeric.
OS Borrelia burgdorferi.
PN WO9512676-A1.
PD 11-MAY-1995.
PF 27-OCT-1994; U12352.
PR 01-NOV-1993; US-148191.
PR 29-APR-1994; US-235836.
PA (ASUY-) ASSOC UNIVERSITIES INC.
PI Dunn JJ, Luft BJ;
DR WPI; 95-215034/28.
DR N-PSDB; Q90717.
PT Chimeric protein comprising 2 or more antigenic Borrelia
PT polypeptide(s) - useful in a vaccine against Lyme borreliosis and in
PT immuno:diagnostic assays
PS Example 1; Fig 15; 200pp; English.
CC The present sequence is the B. burgdorferi strain pTrob, outer
CC surface protein C (OspC-pTrob). Using chemical or enzymatic methods,
CC peptide fragments of OspC-pTrob were prepd., and analysed by western
CC blot to assess their ability to bind different anti-OspC monoclonal
CC antibodies. The information obtd. was used to locate antigenic
CC domains in OspC-pTrob, the epitopes of which were mapped with the
CC aid of site directed mutagenesis. Identical analyses were performed
CC on a selection of Osp purified from a variety of B. burgdorferi
CC strains, the results from which were utilised in the prepn. of a
CC pool of antigenic Borrelia polypeptides, and corresponding
CC polynucleotides. Chimaeric proteins comprising 2 or more antigenic
CC Borrelia polypeptides, that do not naturally occur in the same
CC protein, can be used in the treatment and diagnosis of Borrelia
CC infections, i.e. as a vaccine against Lyme borreliosis, in
CC immunodiagnostic assays to detect anti-Borrelia antibodies or to
CC measure T-cell reactivity.
SQ Sequence 207 AA;

Query Match 54.9%; Score 45; DB 1; Length 233;
Best Local Similarity 55.6%; Pred. No. 9.22e+00;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 67 TVISKKITD 75
:||:||::|
Qy 4 AVITKKVAD 12

RESULT 29
ID R75728 standard; Protein; 209 AA.
AC R75728;
DT 31-JUL-1996 (first entry)
DE B. burgdorferi strain K48 outer surface protein C (OspC-K48).
KW Strain K48; outer surface protein; OspC; antigenic domain;
KW chimaeric protein; treatment; diagnosis; infection; vaccine;
KW Lyme borreliosis; immunodiagnostic assay; antibody;
KW T-cell reactivity; chimeric.
OS Borrelia burgdorferi.
PN WO9512676-A1.
PD 11-MAY-1995.
PF 27-OCT-1994; U12352.
PR 01-NOV-1993; US-148191.
PR 29-APR-1994; US-235836.
PA (ASUY-) ASSOC UNIVERSITIES INC.
PI Dunn JJ, Luft BJ;
DR WPI; 95-215034/28.
DR N-PSDB; Q90715.
PT Chimeric protein comprising 2 or more antigenic Borrelia
PT polypeptide(s) - useful in a vaccine against Lyme borreliosis and in
PT immuno:diagnostic assays
PS Example 1; Fig 13; 200pp; English.
CC The present sequence is the B. burgdorferi strain K48, outer
CC surface protein C (OspC-K48). Using chemical or enzymatic methods,
CC peptide fragments of OspC-K48 were prepd., and analysed by western
CC blot to assess their ability to bind different anti-OspC monoclonal
CC antibodies. The information obtd. was used to locate antigenic
CC domains in OspC-K48, the epitopes of which were mapped with the
CC aid of site directed mutagenesis. Identical analyses were performed
CC on a selection of Osp purified from a variety of B. burgdorferi
CC strains, the results from which were utilised in the prepn. of a
CC pool of antigenic Borrelia polypeptides, and corresponding
CC polynucleotides. Chimaeric proteins comprising 2 or more antigenic
CC Borrelia polypeptides, that do not naturally occur in the same
CC protein, can be used in the treatment and diagnosis of Borrelia
CC infections, i.e. as a vaccine against Lyme borreliosis, in
CC immunodiagnostic assays to detect anti-Borrelia antibodies or to
CC measure T-cell reactivity.
SQ Sequence 209 AA;

Query Match 54.9%; Score 45; DB 1; Length 235;
Best Local Similarity 55.6%; Pred. No. 9.22e+00;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 67 TVISKKITD 75
:||:||::|
Qy 4 AVITKKVAD 12

RESULT 30
ID W02719 standard; peptide; 277 AA.
AC W02719;
DT 13-NOV-1996 (first entry)
DE G-protein coupled odorant receptor I14.
KW G-protein coupled receptor; ligand binding assay; transmembrane domain;
KW schizophrenia; dopamine; cAMP; adenosine; thrombin; adrenergic; opsin;
KW muscarinic acetylcholine; endothelin; bombesin; endocrine; rhodopsin;
KW odorant; cytomegalovirus; serotonergic.
OS Synthetic.
PN US5508384-A.
PD 16-APR-1996.
PF 10-SEP-1992; 943236.
PR 10-SEP-1992; US-943236.
PR 09-SEP-1993; US-118270.
PA (UYNY) UNIV NEW YORK STATE.
PI Murphy RB, Schuster DI;
DR WPI; 96-208785/21.
PT New dopamine receptor peptide - useful as antipsychotic agent, e.g.
PT for treating schizophrenia
PS Disclosure; Column 159-160; 184pp; English.
CC Proteins W02657-W02730 represent a range of G-protein coupled receptor
CC (GPR) proteins selected from cAMP, adenosine, muscarinic acetylcholine,
CC adrenergic, thrombin, endothelin, bombesin, endocrine, rhodopsin, opsin,
CC odorant, cytomegaloviral and other GPR proteins. The receptor proteins
CC were used to design polypeptides, pref. based on the transmembrane
CC domains, for use in G-protein coupled receptor ligand binding assays.
CC The polypeptide fragments retain biological activity such as binding a
CC GPR ligand or modulating GPR ligand binding to a GPR (see W02747-W02999
CC for examples of polypeptide fragments). The polypeptide fragments can
CC be used in compositions for treating subjects suffering from a pathology
CC related to a GPR abnormality e.g. a psychotic disorder such as
CC schizophrenia.
SQ Sequence 277 AA;

Query Match 54.9%; Score 45; DB 1; Length 303;
Best Local Similarity 45.5%; Pred. No. 9.22e+00;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 292 LIRVICTKKIS 302
|:| : |||::
Qy 1 LFRAVITKKVA 11

RESULT 31
ID R48747 standard; Protein; 277 AA.
AC R48747;
DT 07-JUN-1996 (first entry)
DE G-protein coupled odorant receptor I14 protein.
KW G-protein coupled receptor; ligand binding assay; transmembrane domain;
KW psychotic disorder; schizophrenia; dopamine; cAMP; adenosine; thrombin;
KW muscarinic acetylcholine; adrenergic; endothelin; bombesin; endocrine;
KW rhodopsin; opsin; odorant; cytomegalovirus.
OS Synthetic.

```
PN   WO9405695-A1.
PD   17-MAR-1994.
PF   09-SEP-1993; U08528.
PR   10-SEP-1992; US-943236.
PA   (UYNY ) UNIV NEW YORK STATE.
PI   Murphy RB,  Schuster DI;
DR   WPI; 94-101120/12.
PT   Polypeptides of G-coupled receptor proteins (GPRs) - useful for
PT   binding GPR ligands or modulating GPR binding
PS   Disclosure; Page 122-123; 160pp; English.
CC   Proteins R48685-R48758 represent a range of G-protein coupled receptor
CC   proteins selected from cAMP, adenosine, muscarinic acetylcholine,
CC   adrenergic, thrombin, endothelin, bombesin, endocrine, rhodopsin, opsin,
CC   odorant, cytomegaloviral and other G-protein coupled receptors.  The
CC   receptor proteins were used to design polypeptides, pref. based on the
CC   transmembrane domains, for use in G-protein coupled receptor ligand
CC   binding assays.  The polypeptide fragments retain biological activity
CC   such as binding a GPR ligand or modulating GPR ligand binding to a GPR
CC   (see R48759-R48758, R50569-R50807 and R89189-R89195 for examples of
CC   polypeptide fragments).  The polypeptide fragments can be used in
CC   compositions for treating subjects suffering from a pathology related to
CC   a GPR abnormality e.g. a psychotic disorder such as schizophrenia.
SQ   Sequence   277 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 303;
  Best Local Similarity  45.5%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative     4;  Mismatches    2;  Indels    0;  Gaps    0;

Db     292 LIRVICTKKIS 302
           |:| : |||::
Qy       1 LFRAVITKKVA 11


RESULT   32
ID   R27875 standard; Protein; 312 AA.
AC   R27875;
DT   15-MAR-1993  (first entry)
DE   Odorant receptor clone I14.
KW   Odorant receptor; insect; vertebrate; fish; mammal; neurotransmitter;
KW   hormone; G-protein; surface receptor; olfactory epithelium; PCR;
KW   Sprague-Dawley rat; amplify; primer; polymerase chain reaction;
KW   multigene family; ligand binding domain.
OS   Ratus ratus.
PN   WO9217585-A.
PD   15-OCT-1992.
PF   06-APR-1992; U02741.
PR   05-APR-1991; US-681880.
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
PI   Axel R, Buck LB;
DR   WPI; 92-366257/44.
DR   N-PSDB; Q29863.
PT   Nucleic acid encoding an odorant receptor - can be used to
PT   control insect populations or for detecting odours e.g. alcohol,
PT   explosives, natural gas etc.
PS   Claim 44; Fig 17; 195pp; English.
CC   The sequences given in R27867-89 are encoded by odorant receptor
CC   clones derived from an insect, a vertebrate, a fish or a mammal.
CC   These clones form a family of neurotransmitters and hormone receptors
CC   which transduce intracellular signals by activation of specific G-
CC   proteins.  Each of these receptors is a member of a superfamily of
CC   surface receptors which traverse the membrane seven times.  These
CC   clones are only expressed in the olfactory epithelium.  These clones
CC   were isolated using probes derived from RNA prepared from the
CC   olfactory epithelia of Sprague-Dawley rats.  Isolated cDNA's were
CC   amplified using primers which correspond to transmembrane domain 2
CC   and 7.  PCR products of the appropriate size were isolated and
CC   sequenced.  The deduced protein sequences of these cDNA's defined a
CC   new multigene family which shared sequence and structural properties
CC   with the superfamily of neurotransmitter and hormone receptors which
CC   traverse the membrane seven times.  This novel family, however
CC   exhibits features different from any other member of the superfamily
CC   identified so far.  There is a striking divergence within the third,
CC   fourth and fifth transmembrane domains between the olfactory proteins.
CC   This divergence in the potential ligand binding domain is consistent
CC   with the idea that the family of molecules cloned is capable of
CC   asssociating with a large number of odorant of diverse molecular
CC   structure.
SQ   Sequence   312 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 338;
  Best Local Similarity  45.5%;  Pred. No. 9.22e+00;
  Matches     5;  Conservative     4;  Mismatches    2;  Indels    0;  Gaps    0;

Db     327 LIRVICTKKIS 337
           |:| : |||::
Qy       1 LFRAVITKKVA 11


RESULT   33
ID   R62042 standard; Protein; 434 AA.
AC   R62042;
DT   06-JUL-1995  (first entry)
DE   Dialkylglycine decarboxylase.
KW   Dialkylglycine decarboxylase; dgdR; repressor; repression;
KW   regulation; gene expression.
OS   Pseudomonas cepacia.
PN   US5356796-A.
PD   18-OCT-1994.
PF   30-MAR-1990; 501814.
PR   30-MAR-1990; US-501814.
PR   28-SEP-1992; US-952817.
PA   (UYAL-) UNIV ALASKA.
PI   Keller JW;
DR   WPI; 94-332342/41.
DR   N-PSDB; Q72718.
PT   2,2-di:alkyl:glycine decarboxylase (DD) repressor protein - and
PT   vectors comprising its coding sequence and operators, useful for
PT   prepn. of Pseudomonas cepacia DD and for regulating gene
PT   expression
PS   Disclosure; Figure 3; 52pp; English.
CC   E.coli may be transformed with vectors comprising the gene encoding
CC   the 2,2-dialkylglycine decarboxylase sequence (the gene is induced
CC   by the presence of dialkylglycines), with the specific aim of
CC   producing and isolating the enzyme.  A repressor gene (Q72717) is
CC   useful for regulating expression of 2,2-dialkylglycine carboxylase.
CC   The use of the repressor gene and its product allows the rate of
CC   2,2-dialkylglycine decarboxylase production to be controlled.
SQ   Sequence   434 AA;

  Query Match           54.9%;  Score 45;  DB 1;  Length 460;
  Best Local Similarity  33.3%;  Pred. No. 9.22e+00;
  Matches     4;  Conservative     5;  Mismatches    3;  Indels    0;  Gaps    0;

Db     104 LFSGIVSRPVVD 115
           || ::::: | |
Qy       1 LFRAVITKKVAD 12


RESULT   34
ID   R36724 standard; Protein; 434 AA.
AC   R36724;
DT   03-SEP-1993  (first entry)
DE   2,2-dialkylglycine decarboxylase of Pseudomonas cepacia.
KW   Pyridoxyl 5'-phosphate dependent 2,2-dialkylglycine decarboxylase;
KW   pyruvate; EC 4.1.1.64; soil bacterium; decarboxylation;
KW   aminotransferase; repression-induction system.
OS   Pseudomonas cepacia.
FH   Key             Location/Qualifiers
FT   Peptide         2. .20
FT   /note= "sequenced directly"
FT   Peptide         261. .276
FT   /note= "sequenced directly"
PN   US5210025-A.
PD   11-MAY-1993.
PF   30-MAR-1990; 501814.
PR   30-MAR-1990; US-501814.
```

PA (UYAL-) UNIV ALASKA.
PI Keller JW;
DR WPI; 93-166958/20.
DR N-PSDB; Q41259.
PT Prepn. of 2,2-di:alkyl:glycine decarboxylase of Pseudomonas
PT cepacia - in E.coli cells transformed with vectors regulated by
PT repressor protein
PS Disclosure; Fig 3; 23pp; English.
CC A 3969 bp PstI-PstI fragment of P.cepacia DNA containing the gene
CC for 2,2-dialkylglycine decarboxylase was cloned in E.coli. The
CC C-terminus of the deduced amino acid of the decarboxylase is
CC homologous with the C-terminus of the mammalian ornithine amino-
CC transferase and the active site is similar to several other
CC aminotransferases. No homologies with known decarboxylase sequences
CC could be found. Expression of the decarboxylase is controlled by a
CC 687-nucleotide sequence upstream of and diverging from the
CC structural gene. Expression is induced by S-isovaline,
CC 2-methylalanine and D-2-aminobutanoic acid, but not by glycine,
CC D- or L-alanine, L-2-aminobutanoic acid, R-isovaline, or other
CC alkyl amino acids. See R36725 for repressor sequence.
SQ Sequence 434 AA;

Query Match 54.9%; Score 45; DB 1; Length 460;
Best Local Similarity 33.3%; Pred. No. 9.22e+00;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 104 LFSGIVSRPVVD 115
|| ::::: | |
Qy 1 LFRAVITKKVAD 12

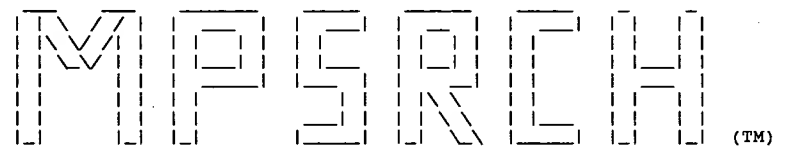
RESULT 35
ID W20991 standard; protein; 614 AA.
AC W20991;
DT 21-JUL-1997 (first entry)
DE H. pylori inner membrane protein, hp3e11168orf29.
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
OS Helicobacter pylori.
PN WO9640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR) ASTRA AB.
PI Berglindh OT, Smith D, Mellgaerd BL;
DR WPI; 97-052306/05.
DR N-PSDB; T68244.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 56; Page 1379-1380; 1481pp; English.
CC The present sequence is a H. pylori inner membrane protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 614 AA;

Query Match 54.9%; Score 45; DB 1; Length 640;
Best Local Similarity 50.0%; Pred. No. 9.22e+00;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 571 FREIIRKEVS 580
|| :| | |:
Qy 2 FRAVITKKVA 11

Search completed: Tue Apr 7 08:41:11 1998
Job time : 7 secs.

****



****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:44:52 1998; MasPar time 2.51 Seconds
26.281 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-3
Description: (1-12) from 5541104.pep
Perfect Score: 82
Sequence: 1 LFRAVITKKVAD 12

Scoring table: PAM 150
Gap 15

Searched: 60183 seqs, 5492030 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: a-issued
1:backl 2:51 3:52 4:53 5:54 6:55 7:56 8:57 9:PCT90
10:PCT91 11:PCT92 12:PCT93 13:PCT94 14:PCT95 15:PCT96

Statistics: Mean 16.538; Variance 48.719; scale 0.339

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	82	100.0	12	6	US-08-190-	Sequence 3, Applicatio	4.26e-03
2	45	54.9	277	12	PCT-US93-0	Sequence 68, Applicati	8.88e+01
3	45	54.9	277	6	US-08-118-	Sequence 68, Applicati	8.88e+01
4	45	54.9	434	4	US-07-952-	Sequence 14, Applicati	8.88e+01
5	45	54.9	434	4	US-07-952-	Sequence 9, Applicatio	8.88e+01
6	45	54.9	434	1	5210025-2	Patent No. 5210025.	8.88e+01
7	45	54.9	434	1	5210025-7	Patent No. 5210025.	8.88e+01
8	43	52.4	121	13	PCT-US94-0	Sequence 22, Applicati	1.44e+02
9	43	52.4	121	6	US-08-180-	Sequence 22, Applicati	1.44e+02
10	43	52.4	121	7	US-08-385-	Sequence 22, Applicati	1.44e+02
11	43	52.4	212	5	US-08-142-	Sequence 5, Applicatio	1.44e+02
12	43	52.4	295	6	US-08-118-	Sequence 79, Applicati	1.44e+02
13	43	52.4	295	12	PCT-US93-0	Sequence 79, Applicati	1.44e+02
14	43	52.4	317	13	PCT-US94-0	Sequence 17, Applicati	1.44e+02
15	43	52.4	317	7	US-08-385-	Sequence 17, Applicati	1.44e+02
16	43	52.4	317	6	US-08-180-	Sequence 17, Applicati	1.44e+02
17	43	52.4	788	5	US-08-194-	Sequence 12, Applicati	1.44e+02
18	43	52.4	1456	6	US-08-026-	Sequence 8, Applicatio	1.44e+02
19	43	52.4	1482	6	US-08-026-	Sequence 2, Applicatio	1.44e+02
20	42	51.2	76	6	US-08-250-	Sequence 1, Applicatio	1.83e+02
21	42	51.2	76	7	US-08-235-	Sequence 1, Applicatio	1.83e+02
22	42	51.2	76	5	US-07-956-	Sequence 1, Applicatio	1.83e+02
23	42	51.2	78	14	PCT-US95-0	Sequence 1, Applicatio	1.83e+02
24	42	51.2	78	7	US-08-330-	Sequence 12, Applicati	1.83e+02
25	42	51.2	99	15	PCT-US96-1	Sequence 5, Applicatio	1.83e+02
26	42	51.2	99	1	5212073-2	Patent No. 5212073.	1.83e+02
27	42	51.2	99	7	US-08-347-	Sequence 8, Applicatio	1.83e+02
28	42	51.2	99	6	US-08-482-	Sequence 35, Applicati	1.83e+02
29	42	51.2	99	7	US-08-480-	Sequence 19, Applicati	1.83e+02
30	42	51.2	99	6	US-08-127-	Sequence 35, Applicati	1.83e+02
31	42	51.2	100	12	PCT-US93-0	Sequence 8, Applicatio	1.83e+02
32	42	51.2	100	4	US-07-820-	Sequence 8, Applicatio	1.83e+02
33	42	51.2	107	14	PCT-US95-1	Sequence 17, Applicati	1.83e+02
34	42	51.2	107	8	US-08-369-	Sequence 17, Applicati	1.83e+02
35	42	51.2	200	13	PCT-US94-1	Sequence 4, Applicatio	1.83e+02
36	42	51.2	220	12	PCT-US93-0	Sequence 4, Applicatio	1.83e+02
37	42	51.2	220	4	US-07-820-	Sequence 4, Applicatio	1.83e+02
38	42	51.2	400	7	US-08-454-	Sequence 6, Applicatio	1.83e+02
39	42	51.2	400	7	US-08-447-	Sequence 6, Applicatio	1.83e+02
40	42	51.2	452	13	PCT-US94-0	Sequence 22, Applicati	1.83e+02
41	42	51.2	847	8	US-08-781-	Sequence 2, Applicatio	1.83e+02
42	42	51.2	847	6	US-08-276-	Sequence 2, Applicatio	1.83e+02
43	42	51.2	859	14	PCT-US95-0	Sequence 2, Applicatio	1.83e+02
44	42	51.2	859	7	US-08-395-	Sequence 2, Applicatio	1.83e+02
45	42	51.2	1121	7	US-08-487-	Sequence 2, Applicatio	1.83e+02
46	42	51.2	1121	5	US-08-005-	Sequence 2, Applicatio	1.83e+02
47	42	51.2	1121	3	US-07-789-	Sequence 2, Applicatio	1.83e+02
48	42	51.2	1220	6	US-08-158-	Sequence 43, Applicati	1.83e+02
49	42	51.2	1289	11	PCT-US92-0	Sequence 4, Applicatio	1.83e+02
50	42	51.2	1289	1	5281530-3	Patent No. 5281530.	1.83e+02
51	42	51.2	1289	1	5426049-4	Patent No. 5426049.	1.83e+02
52	42	51.2	1289	6	US-08-158-	Sequence 4, Applicatio	1.83e+02
53	42	51.2	1289	7	US-08-304-	Sequence 4, Applicatio	1.83e+02
54	42	51.2	1289	3	US-07-675-	Sequence 4, Applicatio	1.83e+02
55	42	51.2	1289	4	US-08-063-	Sequence 4, Applicatio	1.83e+02
56	42	51.2	1289	3	US-07-876-	Sequence 4, Applicatio	1.83e+02
57	42	51.2	1298	14	PCT-US95-0	Sequence 33, Applicati	1.83e+02
58	42	51.2	1298	7	US-08-222-	Sequence 33, Applicati	1.83e+02
59	42	51.2	1334	1	5476657-1	Patent No. 5476657.	1.83e+02
60	42	51.2	1385	11	PCT-US92-0	Sequence 2, Applicatio	1.83e+02
61	42	51.2	1385	4	US-08-063-	Sequence 2, Applicatio	1.83e+02
62	42	51.2	1385	6	US-08-158-	Sequence 2, Applicatio	1.83e+02
63	42	51.2	1385	3	US-07-876-	Sequence 2, Applicatio	1.83e+02
64	42	51.2	1385	1	5426049-1	Patent No. 5426049.	1.83e+02
65	42	51.2	1385	1	5281530-1	Patent No. 5281530.	1.83e+02
66	42	51.2	1385	7	US-08-304-	Sequence 2, Applicatio	1.83e+02
67	42	51.2	1385	3	US-07-675-	Sequence 2, Applicatio	1.83e+02
68	42	51.2	1403	6	US-07-908-	Sequence 3, Applicatio	1.83e+02
69	41	50.0	213	12	PCT-US93-0	Sequence 2, Applicatio	2.32e+02
70	41	50.0	213	8	US-08-459-	Sequence 2, Applicatio	2.32e+02
71	41	50.0	213	7	US-08-459-	Sequence 2, Applicatio	2.32e+02
72	41	50.0	213	7	US-08-287-	Sequence 2, Applicatio	2.32e+02
73	41	50.0	213	7	US-08-460-	Sequence 2, Applicatio	2.32e+02
74	41	50.0	441	7	US-08-403-	Sequence 10, Applicati	2.32e+02
75	41	50.0	604	12	PCT-US93-0	Sequence 2, Applicatio	2.32e+02
76	41	50.0	723	1	5200183-4	Patent No. 5200183.	2.32e+02
77	41	50.0	1023	7	US-08-198-	Sequence 4, Applicatio	2.32e+02
78	41	50.0	1162	14	PCT-US95-1	Sequence 15, Applicati	2.32e+02
79	41	50.0	2165	14	PCT-US95-1	Sequence 2, Applicatio	2.32e+02
80	41	50.0	2368	7	US-08-198-	Sequence 15, Applicati	2.32e+02
81	41	50.0	3080	1	5223423-4	Patent No. 5223423.	2.32e+02
82	40	48.8	37	6	US-08-190-	Sequence 206, Applicat	2.94e+02
83	40	48.8	181	6	US-08-312-	Sequence 28, Applicati	2.94e+02
84	40	48.8	181	13	PCT-US94-0	Sequence 28, Applicati	2.94e+02
85	40	48.8	181	4	US-08-049-	Sequence 28, Applicati	2.94e+02
86	40	48.8	197	6	US-08-456-	Sequence 48, Applicati	2.94e+02
87	40	48.8	218	5	US-08-412-	Sequence 4, Applicatio	2.94e+02
88	40	48.8	267	1	5204252-2	Patent No. 5204252.	2.94e+02
89	40	48.8	268	1	5204252-4	Patent No. 5204252.	2.94e+02
90	40	48.8	312	12	PCT-US93-0	Sequence 38, Applicati	2.94e+02
91	40	48.8	312	6	US-08-118-	Sequence 38, Applicati	2.94e+02
92	40	48.8	358	6	US-08-409-	Sequence 5, Applicatio	2.94e+02
93	40	48.8	481	8	US-08-655-	Sequence 2, Applicatio	2.94e+02
94	40	48.8	576	6	US-08-190-	Sequence 56, Applicati	2.94e+02
95	40	48.8	778	12	PCT-US93-0	Sequence 3, Applicatio	2.94e+02

96	40	48.8	778	12	PCT-US93-0	Sequence 4, Applicatio	2.94e+02
97	40	48.8	1513	12	PCT-US93-0	Sequence 2, Applicatio	2.94e+02
98	40	48.8	2616	1	5206163-3	Patent No. 5206163.	2.94e+02
99	39	47.6	235	13	PCT-US94-0	Sequence 13, Applicati	3.71e+02
100	39	47.6	1480	12	PCT-US93-1	Sequence 2, Applicatio	3.71e+02

ALIGNMENTS

```
RESULT    1
ID   US-08-190-411A-3     STANDARD;     PRT;   12 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 3, Application US/08190411A.
XX
CC    Sequence 3, Application US/08190411A
CC    Patent No. 5541104
CC     GENERAL INFORMATION:
CC       APPLICANT:  Chen, Yao-Tseng; Stockert, Elisabeth;
CC       APPLICANT:  Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.;
CC       APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry;
CC       APPLICANT:  Old, Lloyd J.
CC       TITLE OF INVENTION:  MONOCLONAL ANTIBODIES WHICH BIND TO
CC       TITLE OF INVENTION:  TUMOR REJECTION ANTIGEN PRECURSOR MAGE-1, RECOMBI
  NANT MAGE-1,
CC       TITLE OF INVENTION:  AND MAGE-1 DERIVED IMMUNOGENIC PEPTIDES
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/190,411A
CC         FILING DATE:  01-FEBRUARY-1994
CC         CLASSIFICATION:  436
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  037,230
CC         FILING DATE:  26-MARCH-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/04354
CC         FILING DATE:  22-MAY-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/807,043
CC         FILING DATE:  12-DECEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: *07/764,364
CC         FILING DATE:  23-SEPTEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/728,838
CC         APPLICATION NUMBER:  9-JULY-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/705,702
CC         FILING DATE:  23-MAY-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5541104man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5354
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  3:
CC       SEQUENCE CHARACTERISTICS:
CC       * LENGTH:  12 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   12 AA; 1361 MW; 791 CN;

  Query Match           100.0%;  Score 82;  DB 6;  Length 12;
  Best Local Similarity 100.0%;  Pred. No. 4.26e-03;
  Matches    12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 LFRAVITKKVAD 12
           ||||||||||||
Qy       1 LFRAVITKKVAD 12


RESULT    2
ID   PCT-US93-08528-68     STANDARD;     PRT;   277 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 68, Application PC/TUS9308528.
XX
CC    Sequence 68, Application PC/TUS9308528
CC     GENERAL INFORMATION:
CC       APPLICANT:  New York University
CC       TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
CC       TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
CC       NUMBER OF SEQUENCES:  348
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  BROWDY AND NEIMARK
CC         STREET:  419 Seventh Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  USA
CC         ZIP:  20004
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US93/08528
CC         FILING DATE:  09-SEP-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/943,236
CC         FILING DATE:  10-SEP-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Townsend, Kevin G.
CC         REGISTRATION NUMBER:  34,033
CC         REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-628-5197
CC         TELEFAX:  202-737-3528
CC         TELEX:  248633
CC     INFORMATION FOR SEQ ID NO:  68:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  277 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   277 AA; 31920 MW; 451947 CN;

  Query Match           54.9%;  Score 45;  DB 12;  Length 277;
  Best Local Similarity 45.5%;  Pred. No. 8.88e+01;
  Matches    5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     266 LIRVICTKKIS 276
           |:| : |||::
Qy       1 LFRAVITKKVA 11
```

```
RESULT    3
ID   US-08-118-270-68      STANDARD;      PRT;    277 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 68, Application US/08118270.
XX
CC    Sequence 68, Application US/08118270
CC    Patent No. 5508384
CC     GENERAL INFORMATION:
CC       APPLICANT:  Murphy, Randall B.
CC       APPLICANT:  Schuster, David I.
CC       TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
CC       TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
CC       NUMBER OF SEQUENCES:  348
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  BROWDY AND NEIMARK
CC         STREET:  419 Seventh Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  USA
CC         ZIP:  20004
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/118,270
CC         FILING DATE:  09-SEP-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/943,236
CC         FILING DATE:  10-SEP-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Townsend, Kevin G.
CC         REGISTRATION NUMBER:  34,033
CC         REFERENCE/DOCKET NUMBER:  MURPHY=2A
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-628-5197
CC         TELEFAX:  202-737-3528
CC         TELEX:  248633
CC      INFORMATION FOR SEQ ID NO:  68:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  277 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   277 AA; 31920 MW; 451947 CN;

  Query Match           54.9%;  Score 45;  DB 6;  Length 277;
  Best Local Similarity  45.5%;  Pred. No. 8.88e+01;
  Matches     5;  Conservative     4;  Mismatches    2;  Indels    0;  Gaps    0;

Db    266 LIRVICTKKIS 276
          |:| : |||::
Qy      1 LFRAVITKKVA 11


RESULT    4
ID   US-07-952-817-14      STANDARD;      PRT;    434 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 14, Application US/07952817.
XX
CC    Sequence 14, Application US/07952817
CC    Patent No. 5356796
CC     GENERAL INFORMATION:
CC       APPLICANT:  Keller, John W.
CC       TITLE OF INVENTION:  A Repressor Protein and Gene for Regulating
CC       TITLE OF INVENTION:  Expression of Polypeptides and Its Use in the Pre
  paration of
CC       TITLE OF INVENTION:  2,2-Dialkylglycine Decarboxylase of Pseudomonas C
  epacia
CC       NUMBER OF SEQUENCES:  30
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
CC         ADDRESSEE:  Dunner
CC         STREET:  1300 I Street, N.W., Suite 700
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  US
CC         ZIP:  20005-3315
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,817
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Meyers, Kenneth J.
CC         REGISTRATION NUMBER:  25,146
CC         REFERENCE/DOCKET NUMBER:  01120.0002-01000
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-408-4000
CC         TELEFAX:  202-408-4400
CC      INFORMATION FOR SEQ ID NO:  14:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  434 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   434 AA; 46545 MW; 886275 CN;

  Query Match           54.9%;  Score 45;  DB 4;  Length 434;
  Best Local Similarity  33.3%;  Pred. No. 8.88e+01;
  Matches     4;  Conservative     5;  Mismatches    3;  Indels    0;  Gaps    0;

Db     78 LFSGIVSRPVVD 89
          || ::::: | |
Qy      1 LFRAVITKKVAD 12


RESULT    5
ID   US-07-952-817-9      STANDARD;      PRT;    434 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 9, Application US/07952817.
XX
CC    Sequence 9, Application US/07952817
CC    Patent No. 5356796
CC     GENERAL INFORMATION:
CC       APPLICANT:  Keller, John W.
CC       TITLE OF INVENTION:  A Repressor Protein and Gene for Regulating
CC       TITLE OF INVENTION:  Expression of Polypeptides and Its Use in the Pre
  paration of
CC       TITLE OF INVENTION:  2,2-Dialkylglycine Decarboxylase of Pseudomonas C
  epacia
CC       NUMBER OF SEQUENCES:  30
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
CC         ADDRESSEE:  Dunner
CC         STREET:  1300 I Street, N.W., Suite 700
CC         CITY:  Washington
CC         STATE:  D.C.
```

```
CC         COUNTRY:  US
CC         ZIP:  20005-3315
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/952,817
CC         FILING DATE:  19920928
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Meyers, Kenneth J.
CC         REGISTRATION NUMBER:  25,146
CC         REFERENCE/DOCKET NUMBER:  01120.0002-01000
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-408-4000
CC         TELEFAX:  202-408-4400
CC     INFORMATION FOR SEQ ID NO:  9:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  434 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   434 AA; 46545 MW; 886275 CN;

  Query Match           54.9%;  Score 45;  DB 4;  Length 434;
  Best Local Similarity  33.3%;  Pred. No. 8.88e+01;
  Matches    4;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db     78 LFSGIVSRPVVD 89
          || ::::: | |
Qy      1 LFRAVITKKVAD 12


RESULT    6
ID   5210025-2     STANDARD;     PRT;   470 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5210025.
XX
CC   Patent No. 5210025
CC       APPLICANT: KELLER, JOHN W.
CC       TITLE OF INVENTION: REPRESSOR PROTEIN GENE FOR REGULATING
CC   EXPRESSION OF POLYPEPTIDES AND ITS USE IN THE PREPARATION OF
CC   2,2-DIALKYGLYCINE DECARBOXYLASE OF PSEUDOMONAS CEPACIA
CC       NUMBER OF SEQUENCES: 18
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/501,814
CC         FILING DATE: 30-MAR-1990
CC   SEQ ID NO:2:
CC         LENGTH: 434
SQ   SEQUENCE   470 AA; 50510 MW; 1154124 CN;

  Query Match           54.9%;  Score 45;  DB 1;  Length 434;
  Best Local Similarity  33.3%;  Pred. No. 8.88e+01;
  Matches    4;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db     78 LFSGIVSRPVVD 89
          || ::::: | |
Qy      1 LFRAVITKKVAD 12


RESULT    7
ID   5210025-7     STANDARD;     PRT;   470 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5210025.
XX
CC   Patent No. 5210025
CC       APPLICANT: KELLER, JOHN W.
CC       TITLE OF INVENTION: REPRESSOR PROTEIN GENE FOR REGULATING
CC   EXPRESSION OF POLYPEPTIDES AND ITS USE IN THE PREPARATION OF
CC   2,2-DIALKYGLYCINE DECARBOXYLASE OF PSEUDOMONAS CEPACIA
CC       NUMBER OF SEQUENCES: 18
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/501,814
CC         FILING DATE: 30-MAR-1990
CC   SEQ ID NO:7:
CC         LENGTH: 434
SQ   SEQUENCE   470 AA; 50542 MW; 1153697 CN;

  Query Match           54.9%;  Score 45;  DB 1;  Length 434;
  Best Local Similarity  33.3%;  Pred. No. 8.88e+01;
  Matches    4;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db     78 LFSGIVSRPVVD 89
          || ::::: | |
Qy      1 LFRAVITKKVAD 12


RESULT    8
ID   PCT-US94-02629-22     STANDARD;     PRT;   121 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 22, Application PC/TUS9402629.
XX
CC    Sequence 22, Application PC/TUS9402629
CC     GENERAL INFORMATION:
CC       APPLICANT:  King, Te-Piao
CC       TITLE OF INVENTION:  CLONING AND RECOMBINANT PRODUCTION OF
CC       TITLE OF INVENTION:  VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND
CC       TITLE OF INVENTION:  HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED
  THEREON
CC       NUMBER OF SEQUENCES:  62
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Klauber & Jackson
CC         STREET:  411 Hackensack Avenue
CC         CITY:  Hackensack
CC         STATE:  New Jersey
CC         COUNTRY:  USA
CC         ZIP:  07601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US94/02629
CC         FILING DATE:  10-MAR-1994
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/180,209
CC         FILING DATE:  11-JAN-1994
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/031,400
CC         FILING DATE:  11-MAR-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Jackson Esq., David A.
CC         REGISTRATION NUMBER:  26,742
CC         REFERENCE/DOCKET NUMBER:  600-1-074 PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  201 487-5800
CC         TELEFAX:  201 343-1684
CC         TELEX:  133521
CC     INFORMATION FOR SEQ ID NO:  22:
CC       SEQUENCE CHARACTERISTICS:
```

```
CC         LENGTH:  121 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   121 AA; 13168 MW; 74343 CN;

  Query Match           52.4%;  Score 43;  DB 13;  Length 121;
  Best Local Similarity  45.5%;  Pred. No. 1.44e+02;
  Matches     5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     14 FIAMIAKKLVE 24
          | |:|:||: :
Qy      2 FRAVITKKVAD 12


RESULT    9
ID   US-08-180-209B-22     STANDARD;     PRT;   121 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 22, Application US/08180209B.
XX
CC    Sequence 22, Application US/08180209B
CC    Patent No. 5593877
CC     GENERAL INFORMATION:
CC       APPLICANT:  King, Te-Piao
CC       TITLE OF INVENTION:  CLONING AND RECOMBINANT PRODUCTION OF
CC       TITLE OF INVENTION:  VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND
CC       TITLE OF INVENTION:  HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED
CC       TITLE OF INVENTION:  THEREON
CC       NUMBER OF SEQUENCES:  62
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Klauber & Jackson
CC         STREET:  411 Hackensack Avenue
CC         CITY:  Hackensack
CC         STATE:  New Jersey
CC         COUNTRY:  USA
CC         ZIP:  07601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/180,209B
CC         FILING DATE:  11-JAN-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/031,400
CC         FILING DATE:  11-MAR-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Jackson Esq., David A.
CC         REGISTRATION NUMBER:  26,742
CC         REFERENCE/DOCKET NUMBER:  600-1-074 CIP
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  201 487-5800
CC         TELEFAX:  201 343-1684
CC         TELEX:  133521
CC     INFORMATION FOR SEQ ID NO:  22:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  121 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   121 AA; 13168 MW; 74343 CN;

  Query Match           52.4%;  Score 43;  DB 6;  Length 121;
  Best Local Similarity  45.5%;  Pred. No. 1.44e+02;
  Matches     5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     14 FIAMIAKKLVE 24
          | |:|:||: :
Qy      2 FRAVITKKVAD 12


RESULT   10
ID   US-08-385-745-22     STANDARD;     PRT;   121 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 22, Application US/08385745.
XX
CC    Sequence 22, Application US/08385745
CC    Patent No. 5612209
CC     GENERAL INFORMATION:
CC       APPLICANT:  King, Te Piao
CC       TITLE OF INVENTION:  Cloning and Recombinant Production of
CC       TITLE OF INVENTION:  Vespid Venom Phospholipases, and Immunological Th
  erapies
CC       TITLE OF INVENTION:  Based Thereon
CC       NUMBER OF SEQUENCES:  27
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pennie & Edmonds
CC         STREET:  1155 Avenue of the Americas
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  U.S.A.
CC         ZIP:  10036-2711
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/385,745
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/031,400
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Misrock, S. Leslie
CC         REGISTRATION NUMBER:  18,872
CC         REFERENCE/DOCKET NUMBER:  3288-020
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  212 790-9090
CC         TELEFAX:  212 869-8864/9741
CC         TELEX:  66141 PENNIE
CC     INFORMATION FOR SEQ ID NO:  22:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  121 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   121 AA; 13168 MW; 74343 CN;

  Query Match           52.4%;  Score 43;  DB 7;  Length 121;
  Best Local Similarity  45.5%;  Pred. No. 1.44e+02;
  Matches     5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     14 FIAMIAKKLVE 24
          | |:|:||: :
Qy      2 FRAVITKKVAD 12


RESULT   11
ID   US-08-142-897-5     STANDARD;     PRT;   212 AA.
XX
AC   xxxxxx
```

```
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application US/08142897.
XX
CC    Sequence 5, Application US/08142897
CC    Patent No. 5447852
CC     GENERAL INFORMATION:
CC       APPLICANT:  Friedman, Jeffrey S.
CC       APPLICANT:  Weissman, Irving L.
CC       TITLE OF INVENTION:  No. 5447852el Cyclophilins, Associating Proteins
CC       TITLE OF INVENTION:  and Uses
CC       NUMBER OF SEQUENCES:  10
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Tracy J. Dunn
CC         STREET:  One Market Plaza, Steuart Tower, Suite 2000
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  94105
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/142,897
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/005,917
CC         FILING DATE:  15-JAN-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/740,375
CC         FILING DATE:  05-AUG-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Dunn, Tracy D.
CC         REGISTRATION NUMBER:  34,587
CC         REFERENCE/DOCKET NUMBER:  5490A-92-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-326-2400
CC         TELEFAX:  415-326-2422
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  212 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   212 AA; 22794 MW; 257599 CN;

  Query Match           52.4%;  Score 43;  DB 5;  Length 212;
  Best Local Similarity  41.7%;  Pred. No. 1.44e+02;
  Matches     5;  Conservative    2;  Mismatches   5;  Indels   0;  Gaps   0;

Db     58 LFGNVVPKTVEN 69
          ||  |: | | :
Qy      1 LFRAVITKKVAD 12


RESULT   12
ID   US-08-118-270-79     STANDARD;     PRT;   295 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 79, Application US/08118270.
XX
CC    Sequence 79, Application US/08118270
CC    Patent No. 5508384
CC     GENERAL INFORMATION:
CC       APPLICANT:  Murphy, Randall B.
CC       APPLICANT:  Schuster, David I.
CC       TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
CC       TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
CC       NUMBER OF SEQUENCES:  348
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  BROWDY AND NEIMARK
CC         STREET:  419 Seventh Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  USA
CC         ZIP:  20004
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/118,270
CC         FILING DATE:  09-SEP-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/943,236
CC         FILING DATE:  10-SEP-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Townsend, Kevin G.
CC         REGISTRATION NUMBER:  34,033
CC         REFERENCE/DOCKET NUMBER:  MURPHY=2A
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-628-5197
CC         TELEFAX:  202-737-3528
CC         TELEX:  248633
CC     INFORMATION FOR SEQ ID NO:  79:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  295 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   295 AA; 34191 MW; 498086 CN;

  Query Match           52.4%;  Score 43;  DB 6;  Length 295;
  Best Local Similarity  75.0%;  Pred. No. 1.44e+02;
  Matches     6;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    133 FMVVITKK 140
          |  |||||
Qy      2 FRAVITKK 9


RESULT   13
ID   PCT-US93-08528-79     STANDARD;     PRT;   295 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 79, Application PC/TUS9308528.
XX
CC    Sequence 79, Application PC/TUS9308528
CC     GENERAL INFORMATION:
CC       APPLICANT:  New York University
CC       TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
CC       TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
CC       NUMBER OF SEQUENCES:  348
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  BROWDY AND NEIMARK
CC         STREET:  419 Seventh Street, N.W., Suite 300
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  USA
CC         ZIP:  20004
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US93/08528
CC         FILING DATE:  09-SEP-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/943,236
CC         FILING DATE:  10-SEP-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Townsend, Kevin G.
CC         REGISTRATION NUMBER:  34,033
CC         REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202-628-5197
CC         TELEFAX:  202-737-3528
CC         TELEX:  248633
CC     INFORMATION FOR SEQ ID NO:  79:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  295 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   295 AA; 34191 MW; 498086 CN;

  Query Match           52.4%;  Score 43;  DB 12;  Length 295;
  Best Local Similarity  75.0%;  Pred. No. 1.44e+02;
  Matches     6;  Conservative     0;  Mismatches   2;  Indels   0;  Gaps   0;

Db    133 FMVVITKK 140
          |  |||||
Qy      2 FRAVITKK 9


RESULT   14
ID   PCT-US94-02629-17     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 17, Application PC/TUS9402629.
XX
CC    Sequence 17, Application PC/TUS9402629
CC     GENERAL INFORMATION:
CC       APPLICANT:  King, Te-Piao
CC       TITLE OF INVENTION:  CLONING AND RECOMBINANT PRODUCTION OF
CC       TITLE OF INVENTION:  VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND
CC       TITLE OF INVENTION:  HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED
  THEREON
CC       NUMBER OF SEQUENCES:  62
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Klauber & Jackson
CC         STREET:  411 Hackensack Avenue
CC         CITY:  Hackensack
CC         STATE:  New Jersey
CC         COUNTRY:  USA
CC         ZIP:  07601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US94/02629
CC         FILING DATE:  10-MAR-1994
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/180,209
CC         FILING DATE:  11-JAN-1994
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/031,400
CC         FILING DATE:  11-MAR-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Jackson Esq., David A.
CC         REGISTRATION NUMBER:  26,742
CC         REFERENCE/DOCKET NUMBER:  600-1-074 PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  201 487-5800
CC         TELEFAX:  201 343-1684
CC         TELEX:  133521
CC     INFORMATION FOR SEQ ID NO:  17:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  317 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   317 AA; 35708 MW; 516933 CN;

  Query Match           52.4%;  Score 43;  DB 13;  Length 317;
  Best Local Similarity  45.5%;  Pred. No. 1.44e+02;
  Matches     5;  Conservative     4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    129 FIAMIAKKLVE 139
          | |:|:||: :
Qy      2 FRAVITKKVAD 12


RESULT   15
ID   US-08-385-745-17     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 17, Application US/08385745.
XX
CC    Sequence 17, Application US/08385745
CC    Patent No. 5612209
CC     GENERAL INFORMATION:
CC       APPLICANT:  King, Te Piao
CC       TITLE OF INVENTION:  Cloning and Recombinant Production of
CC       TITLE OF INVENTION:  Vespid Venom Phospholipases, and Immunological Th
  erapies
CC       TITLE OF INVENTION:  Based Thereon
CC       NUMBER OF SEQUENCES:  27
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pennie & Edmonds
CC         STREET:  1155 Avenue of the Americas
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  U.S.A.
CC         ZIP:  10036-2711
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/385,745
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/031,400
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Misrock, S. Leslie
CC         REGISTRATION NUMBER:  18,872
CC         REFERENCE/DOCKET NUMBER:  3288-020
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  212 790-9090
CC         TELEFAX:  212 869-8864/9741
CC         TELEX:  66141 PENNIE
CC     INFORMATION FOR SEQ ID NO:  17:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  317 amino acids
CC         TYPE:  amino acid
```

```
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   317 AA; 35708 MW; 516933 CN;

  Query Match           52.4%;  Score 43;  DB 7;  Length 317;
  Best Local Similarity  45.5%;  Pred. No. 1.44e+02;
  Matches    5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    129 FIAMIAKKLVE 139
          | |:|:||: :
Qy      2 FRAVITKKVAD 12


RESULT   16
ID   US-08-180-209B-17     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 17, Application US/08180209B.
XX
CC    Sequence 17, Application US/08180209B
CC    Patent No. 5593877
CC     GENERAL INFORMATION:
CC       APPLICANT:  King, Te-Piao
CC       TITLE OF INVENTION:  CLONING AND RECOMBINANT PRODUCTION OF
CC       TITLE OF INVENTION:  VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND
CC       TITLE OF INVENTION:  HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED
CC       TITLE OF INVENTION:  THEREON
CC       NUMBER OF SEQUENCES:  62
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Klauber & Jackson
CC         STREET:  411 Hackensack Avenue
CC         CITY:  Hackensack
CC         STATE:  New Jersey
CC         COUNTRY:  USA
CC         ZIP:  07601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/180,209B
CC         FILING DATE:  11-JAN-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/031,400
CC         FILING DATE:  11-MAR-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Jackson Esq., David A.
CC         REGISTRATION NUMBER:  26,742
CC         REFERENCE/DOCKET NUMBER:  600-1-074 CIP
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  201 487-5800
CC         TELEFAX:  201 343-1684
CC         TELEX:  133521
CC     INFORMATION FOR SEQ ID NO:  17:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  317 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   317 AA; 35708 MW; 516933 CN;

  Query Match           52.4%;  Score 43;  DB 6;  Length 317;
  Best Local Similarity  45.5%;  Pred. No. 1.44e+02;
  Matches    5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    129 FIAMIAKKLVE 139
          | |:|:||: :
Qy      2 FRAVITKKVAD 12


RESULT   17
ID   US-08-194-338-12     STANDARD;     PRT;   788 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 12, Application US/08194338.
XX
CC    Sequence 12, Application US/08194338
CC    Patent No. 5474898
CC     GENERAL INFORMATION:
CC       APPLICANT:  Venter, John C.
CC       APPLICANT:  Fraser, Claire M.
CC       APPLICANT:  McCombie, William R.
CC       TITLE OF INVENTION:  OCTOPAMINE RECEPTOR
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Knobbe, Martens, olson and Bear
CC         STREET:  620 Newport Center Drive, Sixteenth Floor
CC         CITY:  Newport Beach
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  92660
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/194,338
CC         FILING DATE:  08-FEB-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/676,174
CC         FILING DATE:  28-MAR-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Israelsen, Ned A.
CC         REGISTRATION NUMBER:  29,655
CC         REFERENCE/DOCKET NUMBER:  NIH101.001DV1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 235-8550
CC         TELEFAX:  (619) 235-0176
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  788 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       HYPOTHETICAL:  NO
CC       ANTI-SENSE:  NO
CC       FRAGMENT TYPE:  internal
SQ   SEQUENCE   788 AA; 85060 MW; 3068211 CN;

  Query Match           52.4%;  Score 43;  DB 5;  Length 788;
  Best Local Similarity  45.5%;  Pred. No. 1.44e+02;
  Matches    5;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db     21 LFRTVTTSTTT 31
          |||:| |   :
Qy      1 LFRAVITKKVA 11


RESULT   18
ID   US-08-026-138E-8     STANDARD;     PRT;   1456 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
```

```
DE   Sequence 8, Application US/08026138E.
XX
CC    Sequence 8, Application US/08026138E
CC    Patent No. 5502166
CC     GENERAL INFORMATION:
CC       APPLICANT:  Masayoshi MISHINA
CC       TITLE OF INVENTION:  NOVEL PROTEINS AND GENES CODING THE SAME
CC       NUMBER OF SEQUENCES:  19
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Nishiohata Residence 1-107
CC         STREET:  5214, Nishiohata-machi
CC         CITY:  Niigata-shi
CC         STATE:  Niigata-ken
CC         COUNTRY:  JAPAN
CC         ZIP:  951
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 MB storage
CC         COMPUTER:  IBM Compatible
CC         OPERATING SYSTEM:  MS-DOS v.5
CC         SOFTWARE:  Word Perfect 5.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/026,138E
CC         FILING DATE:  26-FEB-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 39563/1992
CC         FILING DATE:  26-FEB-1992
CC         APPLICATION NUMBER:  JP 173155/1992
CC         FILING DATE:  30-JUN-1992
CC         APPLICATION NUMBER:  JP 215017/1992
CC         FILING DATE:  12-AUG-1992
CC         APPLICATION NUMBER:  JP 303878/1992
CC         FILING DATE:  13-NOV-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hamburg, C.Bruce
CC         REGISTRATION NUMBER:  22,389
CC         REFERENCE/DOCKET NUMBER:  F-4551
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 986-2340
CC         TELEFAX:  (212) 953-7733
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1456 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single strand
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       ORIGINAL SOURCE:
CC         ORGANISM:  mouse
CC         TISSUE TYPE:  brain
CC       PUBLICATION INFORMATION:
CC         AUTHORS:  Masayoshi MISHINA
CC         TITLE:  NOVEL PROTEINS AND GENES CODING THE SAME
CC         RELEVANT RESIDUES IN SEQ ID NO:  8: FROM 1 to 4368
SQ   SEQUENCE   1456 AA; 162890 MW; 10978415 CN;

  Query Match           52.4%;  Score 43;  DB 6;  Length 1456;
  Best Local Similarity  62.5%;  Pred. No. 1.44e+02;
  Matches     5;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db   1393 FRALVTNK 1400
          |||::|:|
Qy      2 FRAVITKK 9


RESULT   19
ID   US-08-026-138E-2     STANDARD;     PRT;   1482 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/08026138E.
XX
CC    Sequence 2, Application US/08026138E
CC    Patent No. 5502166
CC     GENERAL INFORMATION:
CC       APPLICANT:  Masayoshi MISHINA
CC       TITLE OF INVENTION:  NOVEL PROTEINS AND GENES CODING THE SAME
CC       NUMBER OF SEQUENCES:  19
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Nishiohata Residence 1-107
CC         STREET:  5214, Nishiohata-machi
CC         CITY:  Niigata-shi
CC         STATE:  Niigata-ken
CC         COUNTRY:  JAPAN
CC         ZIP:  951
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 MB storage
CC         COMPUTER:  IBM Compatible
CC         OPERATING SYSTEM:  MS-DOS v.5
CC         SOFTWARE:  Word Perfect 5.1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/026,138E
CC         FILING DATE:  26-FEB-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  JP 39563/1992
CC         FILING DATE:  26-FEB-1992
CC         APPLICATION NUMBER:  JP 173155/1992
CC         FILING DATE:  30-JUN-1992
CC         APPLICATION NUMBER:  JP 215017/1992
CC         FILING DATE:  12-AUG-1992
CC         APPLICATION NUMBER:  JP 303878/1992
CC         FILING DATE:  13-NOV-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hamburg, C.Bruce
CC         REGISTRATION NUMBER:  22,389
CC         REFERENCE/DOCKET NUMBER:  F-4551
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 986-2340
CC         TELEFAX:  (212) 953-7733
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1482 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single strand
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       ORIGINAL SOURCE:
CC         ORGANISM:  mouse
CC         TISSUE TYPE:  brain
CC       PUBLICATION INFORMATION:
CC         AUTHORS:  Masayoshi MISHINA
CC         TITLE:  NOVEL PROTEINS AND GENES CODING THE SAME
CC         RELEVANT RESIDUES IN SEQ ID NO:  2: FROM 1 to 1482
SQ   SEQUENCE   1482 AA; 165957 MW; 11386804 CN;

  Query Match           52.4%;  Score 43;  DB 6;  Length 1482;
  Best Local Similarity  62.5%;  Pred. No. 1.44e+02;
  Matches     5;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db   1419 FRALVTNK 1426
          |||::|:|
Qy      2 FRAVITKK 9


RESULT   20
ID   US-08-250-958-1     STANDARD;     PRT;   76 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application US/08250958.
XX
CC    Sequence 1, Application US/08250958
CC    Patent No. 5571713
```

```
CC     GENERAL INFORMATION:
CC       APPLICANT:  LYLE, LEON R.
CC       APPLICANT:  KUNKEL, STEVEN L.
CC       APPLICANT:  STRIETER, ROBERT M.
CC       TITLE OF INVENTION:  THERAPEUTIC TREATMENT FOR INHIBITING
CC       TITLE OF INVENTION:  VASCULAR RESTENOSIS
CC       NUMBER OF SEQUENCES:  10
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Rothwell, Figg, Ernst & Kurz
CC         STREET:  Suite 701-E, 555 Thirteenth St., N.W
CC         CITY:  Washington
CC         STATE:  D. C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20004
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/250,958
CC         FILING DATE:  27-MAY-1994
CC         CLASSIFICATION:  514
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/965,678
CC         FILING DATE:  22-OCT-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WALKER, Barbara W.
CC         REGISTRATION NUMBER:  35,400
CC         REFERENCE/DOCKET NUMBER:  2077-206A
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)783-6040
CC         TELEFAX:  (202)783-6031
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  76 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  NO
CC       FRAGMENT TYPE:  N-terminal
SQ   SEQUENCE   76 AA; 8614 MW; 31299 CN;

  Query Match           51.2%;  Score 42;  DB 6;  Length 76;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     42 IFKTIVAKEI 51
           :|:::::| :
Qy      1 LFRAVITKKV 10


RESULT   21
ID   US-08-235-659-1     STANDARD;     PRT;   76 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application US/08235659.
XX
CC    Sequence 1, Application US/08235659
CC    Patent No. 5605671
CC     GENERAL INFORMATION:
CC       APPLICANT:  Lyle, Leon R.
CC       APPLICANT:  Kunkel, Steven L.
CC       APPLICANT:  Strieter, Robert M.
CC       TITLE OF INVENTION:  LABELLED CHEMOKINE MATERIALS AND
CC       TITLE OF INVENTION:  MEDICAL USES THEREOF
CC       NUMBER OF SEQUENCES:  2
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Rothwell, Figg, Ernst & Kurz
CC         STREET:  Suite 701-E, 555 Thirteenth St., N.W
CC         CITY:  Washington
CC         STATE:  D. C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20004
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/235,659
CC         FILING DATE:  29-APR-1994
CC         CLASSIFICATION:  424
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/956,862
CC         FILING DATE:  05-OCT-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/956,863
CC         FILING DATE:  05-OCT-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WALKER, Barbara W.
CC         REGISTRATION NUMBER:  35,400
CC         REFERENCE/DOCKET NUMBER:  2077-205A
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)783-6040
CC         TELEFAX:  (202)783-6031
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  76 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  not relevant
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       FRAGMENT TYPE:  N-terminal
SQ   SEQUENCE   76 AA; 8667 MW; 31314 CN;

  Query Match           51.2%;  Score 42;  DB 7;  Length 76;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     42 IFKTIVAKEI 51
           :|:::::| :
Qy      1 LFRAVITKKV 10


RESULT   22
ID   US-07-956-862A-1     STANDARD;     PRT;   76 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application US/07956862A.
XX
CC    Sequence 1, Application US/07956862A
CC    Patent No. 5413778
CC     GENERAL INFORMATION:
CC       APPLICANT:  KUNKEL, STEVEN L.
CC       APPLICANT:  LYLE, LEON R.
CC       APPLICANT:  STRIETER, ROBERT M.
CC       TITLE OF INVENTION:  LABELLED MONOCYTE CHEMOATTRACTANT
CC       TITLE OF INVENTION:  PROTEIN MATERIAL AND MEDICAL USES
CC       TITLE OF INVENTION:  THEREOF
CC       NUMBER OF SEQUENCES:  1
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Rothwell, Figg, Ernst & Kurz
CC         STREET:  Suite 701-E, 555 Thirteenth St., N.W
CC         CITY:  Washington
CC         STATE:  D. C.
CC         COUNTRY:  U.S.A.
CC         ZIP:  20004
CC       COMPUTER READABLE FORM:
```

```
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/956,862A
CC         FILING DATE:  05-OCT-1992
CC         CLASSIFICATION:  424
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  REPPER, GEORGE R.
CC         REGISTRATION NUMBER:  31,414
CC         REFERENCE/DOCKET NUMBER:  1670-197A
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)783-6040
CC         TELEFAX:  (202)783-6031
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  76 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  NO
CC       FRAGMENT TYPE:  N-terminal
SQ   SEQUENCE   76 AA; 8667 MW; 31314 CN;

  Query Match           51.2%;  Score 42;  DB 5;  Length 76;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     42 IFKTIVAKEI 51
           :|:::::| :
Qy      1 LFRAVITKKV 10


RESULT   23
ID   PCT-US95-00605-1     STANDARD;     PRT;   78 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application PC/TUS9500605.
XX
CC    Sequence 1, Application PC/TUS9500605
CC     GENERAL INFORMATION:
CC       APPLICANT:  Lyle, Leon
CC       APPLICANT:  Thomas-Miller, Beth
CC       TITLE OF INVENTION:  THERAPEUTIC TREATMENT FOR INHIBITING
CC       TITLE OF INVENTION:  VASCULAR RESTENOSIS
CC       NUMBER OF SEQUENCES:  23
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Mallinckrodt Medical, Inc.
CC         STREET:  675 McDonnell Boulevard, P.O. Box 5840
CC         CITY:  St. Louis
CC         STATE:  Missouri
CC         COUNTRY:  USA
CC         ZIP:  63134
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US95/00605
CC         FILING DATE:  13-JAN-1995
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/182,917
CC         FILING DATE:  14-JAN-1994
CC         APPLICATION NUMBER:  US 07/965,678
CC         FILING DATE:  22-OCT-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Vacca, Rita D.
CC         REGISTRATION NUMBER:  33,624
CC         REFERENCE/DOCKET NUMBER:  0783.2
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  314-895-7215
CC         TELEFAX:  314-895-2156
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  78 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  circular
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  NO
CC       ANTI-SENSE:  NO
CC       FEATURE:
CC         NAME/KEY:  Peptide
CC         LOCATION:  5..22
SQ   SEQUENCE   78 AA; 8855 MW; 32943 CN;

  Query Match           51.2%;  Score 42;  DB 14;  Length 78;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     44 IFKTIVAKEI 53
           :|:::::| :
Qy      1 LFRAVITKKV 10


RESULT   24
ID   US-08-330-163-12     STANDARD;     PRT;   78 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 12, Application US/08330163.
XX
CC    Sequence 12, Application US/08330163
CC    Patent No. 5656724
CC     GENERAL INFORMATION:
CC       APPLICANT:  Daly, Thomas J.
CC       APPLICANT:  LaRosa, Gregory J.
CC       TITLE OF INVENTION:  Chemokine-Like Proteins and Methods of
CC       TITLE OF INVENTION:  Use
CC       NUMBER OF SEQUENCES:  46
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Fish & Richardson
CC         STREET:  225 Franklin Street
CC         CITY:  Boston
CC         STATE:  MA
CC         COUNTRY:  U.S.A.
CC         ZIP:  02110-2804
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30B
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/330,163
CC         FILING DATE:  05-AUG-1994
CC         CLASSIFICATION:  530
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Fasse, J. Peter
CC         REGISTRATION NUMBER:  32,983
CC         REFERENCE/DOCKET NUMBER:  00231/080001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (617) 542-5070
CC         TELEFAX:  (617) 542-8906
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  78 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
```

```
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   78 AA; 8953 MW; 32898 CN;

  Query Match           51.2%;  Score 42;  DB 7;  Length 78;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     44 IFKTIVAKEI 53
          :|:::::| :
Qy      1 LFRAVITKKV 10


RESULT   25
ID   PCT-US96-10087-5     STANDARD;     PRT;   99 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 5, Application PC/TUS9610087.
XX
CC    Sequence 5, Application PC/TUS9610087
CC     GENERAL INFORMATION:
CC       APPLICANT:
CC       TITLE OF INVENTION:  Monocyte Chemotactic Protein-4
CC       NUMBER OF SEQUENCES:  6
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US96/10087
CC         FILING DATE:  07-JUN-1996
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/479,126
CC         FILING DATE:  07-JUN-1995
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  99 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   99 AA; 11025 MW; 51738 CN;

  Query Match           51.2%;  Score 42;  DB 15;  Length 99;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     65 IFKTIVAKEI 74
          :|:::::| :
Qy      1 LFRAVITKKV 10


RESULT   26
ID   5212073-2     STANDARD;     PRT;   107 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Patent No. 5212073.
XX
CC   Patent No. 5212073
CC       APPLICANT: ROLLINS, BARRETT;STILES, CHARLES;WONG, GORDON G.
CC       TITLE OF INVENTION: PROCESS FOR PRODUCING HUMAN JE CYTOKINE
CC       NUMBER OF SEQUENCES: 1
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/351,008
CC         FILING DATE:12-MAY-1989
CC ,SEQ ID NO:2:
CC       ' LENGTH: 99
SQ   SEQUENCE   107 AA; 11906 MW; 65752 CN;

  Query Match           51.2%;  Score 42;  DB 1;  Length 99;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     65 IFKTIVAKEI 74
          :|:::::| :
Qy      1 LFRAVITKKV 10


RESULT   27
ID   US-08-347-492B-8     STANDARD;     PRT;   99 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/08347492B.
XX
CC    Sequence 8, Application US/08347492B
CC    Patent No. 5602008
CC     GENERAL INFORMATION:
CC       APPLICANT:  Wilde, Craig G.
CC       APPLICANT:  Hawkins, Phillip R.
CC       APPLICANT:  Bandman, Olga
CC       APPLICANT:  Seilhamer, Jeffrey J.
CC       TITLE OF INVENTION:  EXPRESSED CHEMOKINES, THEIR
CC       TITLE OF INVENTION:  PRODUCTION AND USES
CC       NUMBER OF SEQUENCES:  12
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Incyte Pharmaceuticals, Inc.
CC         STREET:  3174 Porter Drive
CC         CITY:  Palo Alto
CC         STATE:  CA
CC         COUNTRY:  U.S.
CC         ZIP:  94304
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette
CC         COMPUTER:  IBM Compatible
CC         OPERATING SYSTEM:  DOS
CC         SOFTWARE:  FastSEQ Version 1.5
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/347,492B
CC         FILING DATE:  29-NOV-1994
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/303,241
CC         FILING DATE:  07-SEP-1994
CC         APPLICATION NUMBER:  08/320,011
CC         FILING DATE:  05-OCT-1994
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Luther, Barbara J
CC         REGISTRATION NUMBER:  33,954
CC         REFERENCE/DOCKET NUMBER:  PF-0024
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415-855-0555
CC         TELEFAX:  415-852-0195
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  99 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       IMMEDIATE SOURCE:
CC         LIBRARY:  GENBANK
CC         CLONE:  GI 487124
SQ   SEQUENCE   99 AA; 11025 MW; 51738 CN;

  Query Match           51.2%;  Score 42;  DB 7;  Length 99;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;
```

```
Db      65 IFKTIVAKEI 74
           :|:::::| :
Qy       1 LFRAVITKKV 10


RESULT   28
ID   US-08-482-847-35     STANDARD;     PRT;   99 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 35, Application US/08482847.
XX
CC    Sequence 35, Application US/08482847
CC    Patent No. 5556757
CC     GENERAL INFORMATION:
CC       APPLICANT:  VAN ALSTYNE, Diane
CC       APPLICANT:  SHARMA, Lawrence Rajendra
CC       TITLE OF INVENTION:  PEPTIDES REPRESENTING EPITOPIC SITES FOR
CC       TITLE OF INVENTION:  BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND
   THEIR
CC       TITLE OF INVENTION:  CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF
CC       NUMBER OF SEQUENCES:  40
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Foley & Lardner
CC         STREET:  3000 K Street, N.W., Suite 500
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  USA
CC         ZIP:  20007-5109
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/482,847
CC         FILING DATE:  07-JUN-1995
CC         CLASSIFICATION:  514
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/127,499
CC         FILING DATE:  28-SEP-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  BENT, Stephen A.
CC         REGISTRATION NUMBER:  29,768
CC         REFERENCE/DOCKET NUMBER:  51916/104/INBI
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)672-5300
CC         TELEFAX:  (202)672-5399
CC         TELEX:  904136
CC      INFORMATION FOR SEQ ID NO:  35:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  99 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:
CC         TOPOLOGY:  unknown
SQ   SEQUENCE   99 AA; 11025 MW; 51738 CN;

  Query Match           51.2%;  Score 42;  DB 6;  Length 99;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db      65 IFKTIVAKEI 74
           :|:::::| :
Qy       1 LFRAVITKKV 10


RESULT   29
ID   US-08-480-449-19     STANDARD;     PRT;   99 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 19, Application US/08480449.
XX
CC    Sequence 19, Application US/08480449
CC    Patent No. 5688927
CC     GENERAL INFORMATION:
CC       APPLICANT:  Godiska, Ronald
CC       APPLICANT:  Gray, Patrick W.
CC       TITLE OF INVENTION:  MACROPHAGE DERIVED CHEMOKINE
CC       NUMBER OF SEQUENCES:  24
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
CC         STREET:  6300 Sears Tower, 233 South Wacker Drive
CC         CITY:  Chicago
CC         STATE:  Illinois
CC         COUNTRY:  United States of America
CC         ZIP:  60606-6402
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/480,449
CC         FILING DATE:
CC         CLASSIFICATION:  530
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Gass, David A.
CC         REGISTRATION NUMBER:  38,153
CC         REFERENCE/DOCKET NUMBER:  27866/32779
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  312/474-6300
CC         TELEFAX:  312/474-0448
CC         TELEX:  25-3856
CC      INFORMATION FOR SEQ ID NO:  19:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  99 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       FEATURE:
CC         NAME/KEY:  misc_feature
CC         OTHER INFORMATION:  "Hu MCP-1"
SQ   SEQUENCE   99 AA; 11025 MW; 51738 CN;

  Query Match           51.2%;  Score 42;  DB 7;  Length 99;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db      65 IFKTIVAKEI 74
           :|:::::| :
Qy       1 LFRAVITKKV 10


RESULT   30
ID   US-08-127-499A-35     STANDARD;     PRT;   99 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 35, Application US/08127499A.
XX
CC    Sequence 35, Application US/08127499A
CC    Patent No. 5510264
CC     GENERAL INFORMATION:
CC       APPLICANT:  VAN ALSTYNE, Diane
CC       APPLICANT:  SHARMA, Lawrence Rajendra
CC       TITLE OF INVENTION:  ANTIBODIES WHICH BIND MENINGITIS RELATED
CC       TITLE OF INVENTION:  HOMOLOGOUS ANTIGENIC SEQUENCES
CC       NUMBER OF SEQUENCES:  40
```

```
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Foley & Lardner
CC         STREET:  3000 K Street, N.W., Suite 500
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  USA
CC         ZIP:  20007-5109
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/127,499A
CC         FILING DATE:  28-SEP-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  BENT, Stephen A.
CC         REGISTRATION NUMBER:  29,768
CC         REFERENCE/DOCKET NUMBER:  51916/102/INBI
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (202)672-5300
CC         TELEFAX:  (202)672-5399
CC         TELEX:  904136
CC     INFORMATION FOR SEQ ID NO:  35:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  99 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:
CC         TOPOLOGY:  unknown
SQ   SEQUENCE   99 AA; 11025 MW; 51738 CN;

  Query Match           51.2%;  Score 42;  DB 6;  Length 99;
  Best Local Similarity  20.0%;  Pred. No. 1.83e+02;
  Matches     2;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     65 IFKTIVAKEI 74
          :|:::::| :
Qy      1 LFRAVITKKV 10


RESULT   31
ID   PCT-US93-00324-8     STANDARD;     PRT;   100 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application PC/TUS9300324.
XX
CC    Sequence 8, Application PC/TUS9300324
CC     GENERAL INFORMATION:
CC       APPLICANT:  Cochran Ph.D., Mark D
CC       APPLICANT:  Junker M.S., David E
CC       TITLE OF INVENTION:  Recombinant Swinepox Virus
CC       NUMBER OF SEQUENCES:  42
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  John P. White
CC         STREET:  30 Rockefeller Plaza
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US93/00324
CC         FILING DATE:  19930113
CC         CLASSIFICATION:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  White, John P
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212)977-9550
CC         TELEFAX:  (212)664-0525
CC         TELEX:  422523
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  100 amino acids
CC         TYPE:  AMINO ACID
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  YES
CC       ANTI-SENSE:  NO
CC       FRAGMENT TYPE:  C-terminal
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Swinepox virus
CC         STRAIN:  Kasza
CC       POSITION IN GENOME:
CC         MAP POSITION:  ~23.2
CC         UNITS:  %G
SQ   SEQUENCE   100 AA; 11740 MW; 54002 CN;

  Query Match           51.2%;  Score 42;  DB 12;  Length 100;
  Best Local Similarity  44.4%;  Pred. No. 1.83e+02;
  Matches     4;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db     25 LYRILIVKR 33
          |:| :| |:
Qy      1 LFRAVITKK 9


RESULT   32
ID   US-07-820-154A-8     STANDARD;     PRT;   100 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/07820154A.
XX
CC    Sequence 8, Application US/07820154A
CC    Patent No. 5382425
CC     GENERAL INFORMATION:
CC       APPLICANT:  Cochran Ph.D., Mark D
CC       APPLICANT:  Junker M.S., David E
CC       TITLE OF INVENTION:  Recombinant Swinepox Virus
CC       NUMBER OF SEQUENCES:  40
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  John P. White
CC         STREET:  30 Rockefeller Plaza
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10112
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/820,154A
CC         FILING DATE:  19920113
CC         CLASSIFICATION:  424
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  White, John P
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212)977-9550
CC         TELEFAX:  (212)664-0525
CC         TELEX:  422523
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  100 amino acids
CC         TYPE:  AMINO ACID
```

```
CC         STRANDEDNESS:  double
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  YES
CC       ANTI-SENSE:  NO
CC       FRAGMENT TYPE:  C-terminal
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Swinepox virus
CC         STRAIN:  Kasza
CC       POSITION IN GENOME:
CC         MAP POSITION:  -23.2
CC         UNITS:  %G
SQ   SEQUENCE   100 AA; 11740 MW; 54002 CN;

  Query Match           51.2%;  Score 42;  DB 4;  Length 100;
  Best Local Similarity  44.4%;  Pred. No. 1.83e+02;
  Matches     4;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db     25 LYRILIVKR 33
          |:| :| |:
Qy      1 LFRAVITKK 9


RESULT   33
ID   PCT-US95-17025-17     STANDARD;     PRT;   107 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 17, Application PC/TUS9517025.
XX
CC    Sequence 17, Application PC/TUS9517025
CC     GENERAL INFORMATION:
CC       APPLICANT:  James E. Darnell, Jr.
CC       APPLICANT:  Zilong Wen
CC       APPLICANT:  Curt M. Horvath
CC       APPLICANT:  Zhong Zhong
CC       TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
CC       TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT)
  PROTEINS
CC       NUMBER OF SEQUENCES:  39
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Klauber & Jackson
CC         STREET:  411 Hackensack Avenue
CC         CITY:  Hackensack
CC         STATE:  New Jersey
CC         COUNTRY:  USA
CC         ZIP:  07601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US95/17025
CC         FILING DATE:  28-DEC-1995
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/369,796
CC         FILING DATE:  06-JAN-1995
CC         CLASSIFICATION:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Jackson Esq., David A.
CC         REGISTRATION NUMBER:  26,742
CC         REFERENCE/DOCKET NUMBER:  600-1-116
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  201 487-5800
CC         TELEFAX:  201 343-1684
CC         TELEX:  133521
CC     INFORMATION FOR SEQ ID NO:  17:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  107 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  NO
CC       FRAGMENT TYPE:  internal
SQ   SEQUENCE   107 AA; 11984 MW; 59812 CN;

  Query Match           51.2%;  Score 42;  DB 14;  Length 107;
  Best Local Similarity  40.0%;  Pred. No. 1.83e+02;
  Matches     4;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db      5 LFKNLLLKKI 14
          ||: :: ||:
Qy      1 LFRAVITKKV 10


RESULT   34
ID   US-08-369-796-17     STANDARD;     PRT;   107 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 17, Application US/08369796.
XX
CC    Sequence 17, Application US/08369796
CC    Patent No. 5716622
CC     GENERAL INFORMATION:
CC       APPLICANT:  James E. Darnell, Jr.
CC       APPLICANT:  Zilong Wen
CC       APPLICANT:  Curt M. Horvath
CC       APPLICANT:  Zhong Zhong
CC       TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
CC       TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT)
  PROTEINS
CC       NUMBER OF SEQUENCES:  39
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Klauber & Jackson
CC         STREET:  411 Hackensack Avenue
CC         CITY:  Hackensack
CC         STATE:  New Jersey
CC         COUNTRY:  USA
CC         ZIP:  07601
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/369,796
CC         FILING DATE:  06-JAN-1995
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Jackson Esq., David A.
CC         REGISTRATION NUMBER:  26,742
CC         REFERENCE/DOCKET NUMBER:  600-1-116
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  201 487-5800
CC         TELEFAX:  201 343-1684
CC         TELEX:  133521
CC     INFORMATION FOR SEQ ID NO:  17:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  107 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  NO
CC       FRAGMENT TYPE:  internal
SQ   SEQUENCE   107 AA; 11984 MW; 59812 CN;

  Query Match           51.2%;  Score 42;  DB 8;  Length 107;
```

```
  Best Local Similarity  40.0%;  Pred. No. 1.83e+02;
  Matches     4;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db      5 LFKNLLLKKI 14
            ||: :: ||:
Qy      1 LFRAVITKKV 10


RESULT   35
ID   PCT-US94-14277-4     STANDARD;     PRT;   200 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application PC/TUS9414277.
XX
CC    Sequence 4, Application PC/TUS9414277
CC     GENERAL INFORMATION:
CC       APPLICANT:  Aguet, Michel
CC       APPLICANT:  Bohni, Ruth
CC       APPLICANT:  Hemmi, Silvio
CC       TITLE OF INVENTION:  Receptor Subunit Polypeptides
CC       NUMBER OF SEQUENCES:  8
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Genentech, Inc.
CC         STREET:  460 Point San Bruno Blvd
CC         CITY:  South San Francisco
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  94080
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  patin (Genentech)
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US94/14277
CC         FILING DATE:  07-DEC-1994
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/164596
CC         FILING DATE:  09-DEC-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Love, Richard B.
CC         REGISTRATION NUMBER:  34,659
CC         REFERENCE/DOCKET NUMBER:  866PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415/225-5530
CC         TELEFAX:  415/952-9881
CC         TELEX:  910/371-7168
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  200 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
SQ   SEQUENCE   200 AA; 22525 MW; 227601 CN;

  Query Match           51.2%;  Score 42;  DB 13;  Length 200;
  Best Local Similarity  62.5%;  Pred. No. 1.83e+02;
  Matches     5;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db    175 LFRALLNK 182
          ||||:: |
Qy      1 LFRAVITK 8


Search completed: Tue Apr  7 08:45:06 1998
Job time : 14 secs.
```

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:40:32 1998; MasPar time 4.22 Seconds
121.776 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-3
Description: (1-12) from 5541104.pep
Perfect Score: 82
Sequence: 1 LFRAVITKKVAD 12

Scoring table: PAM 150
Gap 15

Searched: 195121 seqs, 42852602 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: pir55
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 19.923; Variance 56.762; scale 0.351

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	82	100.0	280	2	JC2358	TOIG of: jc2358 check	6.02e-02
2	60	73.2	124	2	I38663	A;Title: Structure, ch	2.29e+01
3	55	67.1	875	2	S75377	A;Cross-references: EM	8.10e+01
4	54	65.9	693	2	JN0673	TOIG of: jn0673 check	1.04e+02
5	54	65.9	701	2	JN0674	TOIG of: jn0674 check	1.04e+02
6	54	65.9	885	2	S59660	TOIG of: s59660 check	1.04e+02
7	53	64.6	236	2	F64964	A;Title: The complete	1.33e+02
8	53	64.6	904	2	G64840	A;Title: The complete	1.33e+02
9	52	63.4	126	2	S64131	TOIG of: s64131 check	1.70e+02
10	52	63.4	317	2	JC2359	A;Accession: PH1298.	1.70e+02
11	52	63.4	317	2	I38661	A;Accession: PH1297.	1.70e+02
12	52	63.4	489	2	S40731	TOIG of: s40731 check	1.70e+02
13	51	62.2	234	2	I38667	A;Title: Structure, ch	2.16e+02
14	51	62.2	736	2	A27477	C;Genetics:.	2.16e+02
15	51	62.2	980	2	S45444	TOIG of: s45444 check	2.16e+02
16	51	62.2	1729	2	S57596	TOIG of: s57596 check	2.16e+02
17	50	61.0	348	2	S18521	A;Accession: S18521.	2.75e+02
18	50	61.0	394	1	YQBZDH	TOIG of: yqbzdh check	2.75e+02
19	50	61.0	690	2	S64646	TOIG of: s64646 check	2.75e+02
20	50	61.0	890	2	F64991	A;Title: The complete	2.75e+02
21	50	61.0	2560	2	I40457	TOIG of: i40457 check	2.75e+02
22	50	61.0	2560	2	B69681	A;Status: preliminary;	2.75e+02
23	49	59.8	224	1	WMBEHG	A;Accession: JQ1494.	3.49e+02
24	49	59.8	236	2	H64425	A;Authors: Borodovsky,	3.49e+02
25	49	59.8	274	2	S18857	TOIG of: s18857 check	3.49e+02
26	49	59.8	274	1	XNECSD	A;Title: The complete	3.49e+02
27	49	59.8	287	2	H64538	A;Authors: Hayes, W.S.	3.49e+02
28	49	59.8	303	2	H64133	A;Authors: Cotton, M.D	3.49e+02
29	49	59.8	405	2	MTTILTSTFR	This is a DE line.	3.49e+02
30	49	59.8	468	2	A54926	TOIG of: a54926 check	3.49e+02
31	49	59.8	528	2	S32593	TOIG of: s32593 check	3.49e+02
32	49	59.8	660	2	S40098	TOIG of: s40098 check	3.49e+02
33	49	59.8	1454	2	S53398	TOIG of: s53398 check	3.49e+02
34	49	59.8	1906	2	S68235	F;1451-1708/Domain: pr	3.49e+02
35	48	58.5	193	2	C64328	A;Authors: Borodovsky,	4.42e+02
36	48	58.5	319	2	C69507	A;Accession: C69507.	4.42e+02
37	48	58.5	320	2	B37767	TOIG of: b37767 check	4.42e+02
38	48	58.5	336	2	C64410	A;Authors: Borodovsky,	4.42e+02
39	48	58.5	458	2	D64708	A;Authors: Hayes, W.S.	4.42e+02
40	48	58.5	516	2	A31270	A;Accession: S05962.	4.42e+02
41	48	58.5	578	2	A44326	TOIG of: a44326 check	4.42e+02
42	48	58.5	580	2	S25010	TOIG of: s25010 check	4.42e+02
43	48	58.5	754	2	S62561	TOIG of: s62561 check	4.42e+02
44	48	58.5	793	2	S77402	A;Accession: S77402.	4.42e+02
45	47	57.3	61	2	B33833	TOIG of: b33833 check	5.59e+02
46	47	57.3	70	2	S55110	TOIG of: s55110 check	5.59e+02
47	47	57.3	80	2	S59636	A;Cross-references: EM	5.59e+02
48	47	57.3	80	2	S59635	A;Cross-references: EM	5.59e+02
49	47	57.3	87	2	F69542	A;Accession: F69542.	5.59e+02
50	47	57.3	149	2	D64076	A;Authors: Cotton, M.D	5.59e+02
51	47	57.3	173	1	CYPGAA	This is a DE line.	5.59e+02
52	47	57.3	173	1	CYMNAA	TOIG of: cymnaa check	5.59e+02
53	47	57.3	173	1	CYLEAA	TOIG of: cyleaa check	5.59e+02
54	47	57.3	173	1	CYRNAA	This is a DE line.	5.59e+02
55	47	57.3	173	1	CYLPAA	TOIG of: cylpaa check	5.59e+02
56	47	57.3	173	1	CYGCAA	TOIG of: cygcaa check	5.59e+02
57	47	57.3	225	1	QQBE42	A;Title: DNA sequence	5.59e+02
58	47	57.3	249	2	S21935	TOIG of: s21935 check	5.59e+02
59	47	57.3	249	2	S09654	TOIG of: s09654 check	5.59e+02
60	47	57.3	268	2	H64435	A;Authors: Borodovsky,	5.59e+02
61	47	57.3	301	2	S45597	TOIG of: s45597 check	5.59e+02
62	47	57.3	309	2	B49878	A;Cross-references: GB	5.59e+02
63	47	57.3	314	2	JC2361	A;Title: Human gene MA	5.59e+02
64	47	57.3	353	2	S34347	A;Accession: S65461.	5.59e+02
65	47	57.3	354	2	S33309	A;Cross-references: EM	5.59e+02
66	47	57.3	355	2	A41534	A;Accession: A41534.	5.59e+02
67	47	57.3	359	1	RGMSQ	F;183/Modified site: A	5.59e+02
68	47	57.3	359	2	S45700	TOIG of: s45700 check	5.59e+02
69	47	57.3	359	2	S45699	TOIG of: s45699 check	5.59e+02
70	47	57.3	359	2	S30359	TOIG of: s30359 check	5.59e+02
71	47	57.3	359	1	RGHUGY	F;183/Modified site: A	5.59e+02
72	47	57.3	359	1	RGMS11	F;183/Modified site: A	5.59e+02
73	47	57.3	360	2	JN0115	TOIG of: jn0115 check	5.59e+02
74	47	57.3	370	2	A41891	TOIG of: a41891 check	5.59e+02
75	47	57.3	399	2	F69034	A;Accession: F69034.	5.59e+02
76	47	57.3	401	1	A31266	TOIG of: a31266 check	5.59e+02
77	47	57.3	468	2	A64601	A;Authors: Hayes, W.S.	5.59e+02
78	47	57.3	628	2	S22396	A;Accession: S22396.	5.59e+02
79	47	57.3	745	2	G01025	TOIG of: g01025 check	5.59e+02
80	47	57.3	774	2	I48609	TOIG of: i48609 check	5.59e+02
81	47	57.3	775	2	S31333	TOIG of: s31333 check	5.59e+02
82	47	57.3	998	2	S41397	TOIG of: s41397 check	5.59e+02
83	47	57.3	1147	2	A41674	F;702-710/Region: prot	5.59e+02
84	47	57.3	1176	2	JN0583	F;731-739/Region: prot	5.59e+02
85	47	57.3	1450	2	S78060	A;Cross-references: EM	5.59e+02
86	46	56.1	175	2	G69833	A;Experimental source:	7.05e+02
87	46	56.1	215	2	B26473	TOIG of: b26473 check	7.05e+02
88	46	56.1	261	2	MFIKWGDRLM	This is a DE line.	7.05e+02
89	46	56.1	273	2	S67622	TOIG of: s67622 check	7.05e+02
90	46	56.1	342	2	MGSLGAILKH	This is a DE line.	7.05e+02
91	46	56.1	357	2	G69290	A;Accession: G69290.	7.05e+02
92	46	56.1	396	1	AJMZRB	A;Accession: B28180.	7.05e+02
93	46	56.1	489	2	B36395	TOIG of: b36395 check	7.05e+02
94	46	56.1	509	2	D64435	A;Authors: Borodovsky,	7.05e+02
95	46	56.1	548	2	S59133	A;Accession: S59133.	7.05e+02
96	46	56.1	732	2	S73089	A;Cross-references: EM	7.05e+02

```
    97     46   56.1    863  2  S38140      TOIG of: s38140  check  7.05e+02
    98     46   56.1    893  2  S46442      F;536-610/Domain: cyto  7.05e+02
    99     46   56.1   1265  2  S57968      TOIG of: s57968  check  7.05e+02
   100     46   56.1   1556  2  S76781      A;Accession: S76781.    7.05e+02


                              ALIGNMENTS

RESULT    1
ID   JC2358     STANDARD;     PRT;   280 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: jc2358  check: 467  from: 1  to: 280.
XX
CC     TOIG of: jc2358  check: 467  from: 1  to: 280
CC
CC    >P1;JC2358
CC    tumor-associated antigen , MAGE-1 - human
CC    C;Species: Homo sapiens (man)
CC    C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Mar-19
  96
CC    C;Accession: JC2358
CC    R;Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
CC    Biochem. Biophys. Res. Commun. 202, 549-555, 1994
CC    A;Title: Cloning and analysis of MAGE-1-related genes.
CC    A;Reference number: JC2358
CC    A;Accession: JC2358
CC    A;Molecule type: mRNA
CC    A;Residues: 1-280 <DIN>
CC    A;Experimental source: melanoma cell line DM150
CC    C;Genetics:
CC    A;Gene: MAGE
CC    F;161-169/Region: HLA-A1 binding #status predicted
SQ   SEQUENCE   280 AA; 30932 MW; 426797 CN;

  Query Match          100.0%;  Score 82;  DB 2;  Length 280;
  Best Local Similarity 100.0%;  Pred. No. 6.02e-02;
  Matches    12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     97 LFRAVITKKVAD 108
           ||||||||||||
Qy      1 LFRAVITKKVAD 12


RESULT    2
ID   I38663     STANDARD;     PRT;   124 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: Structure, chromosomal localization, and expression of 12 genes o
  f the MAGE family.
XX
CC    A;Title: Structure, chromosomal localization, and expression of 12 genes
  of the MAGE family.
CC    A;Reference number: I38659; MUID:95012457
CC    A;Accession: I38663
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-124 <RES>
CC    A;Cross-references: EMBL:U10689; NID:g533518; PID:g533519
CC    A;Note: MAGE-5a antigen
CC    A;Accession: I38664
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-124 <RE2>
CC    A;Cross-references: EMBL:U10690; NID:g533520; PID:g533521
CC    A;Note: MAGE-5b antigen
CC    C;Genetics:
CC    A;Gene: GDB:MAGE5
CC    A;Cross-references: GDB:331120
CC    A;Map position: Xq28-Xq28
CC    A;Introns: #status absent
SQ   SEQUENCE   124 AA; 13015 MW; 83534 CN;

  Query Match          73.2%;  Score 60;  DB 2;  Length 124;
  Best Local Similarity 66.7%;  Pred. No. 2.29e+01;
  Matches    8;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    104 VFRAALSKKVAD 115
          :||| ::|||||
Qy      1 LFRAVITKKVAD 12


RESULT    3
ID   S75377     STANDARD;     PRT;   875 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Cross-references: EMBL:Y08257; NID:g1707772; PID:e283825; PID:g1707782.
XX
CC    A;Cross-references: EMBL:Y08257; NID:g1707772; PID:e283825; PID:g1707782
CC    A;Experimental source: strain P2
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, S
  eptember 1996
CC    C;Genetics:
CC    A;Start codon: TTG
SQ   SEQUENCE   875 AA; 100392 MW; 4004557 CN;

  Query Match          67.1%;  Score 55;  DB 2;  Length 875;
  Best Local Similarity 77.8%;  Pred. No. 8.10e+01;
  Matches    7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    193 LFRNLITKK 201
          ||| :||||
Qy      1 LFRAVITKK 9


RESULT    4
ID   JN0673     STANDARD;     PRT;   693 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: jn0673  check: 7105  from: 1  to: 693.
XX
CC     TOIG of: jn0673  check: 7105  from: 1  to: 693
CC
CC    >P1;JN0673
CC    ubiquitin-like fusion protein Anla - African clawed frog
CC    C;Species: Xenopus laevis (African clawed frog)
CC    C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 08-Sep-19
  97
CC    C;Accession: JN0673
CC    R;Linnen, J.M.; Bailey, C.P.; Weeks, D.L.
CC    Gene 128, 181-188, 1993
CC    A;Title: Two related localized mRNAs from Xenopus laevis encode ubiquitin
  -like fusion proteins.
CC    A;Reference number: JN0673
CC    A;Accession: JN0673
CC    A;Status: nucleic acid sequence not shown
CC    A;Molecule type: mRNA
CC    A;Residues: 1-693 <LIN>
CC    C;Genetics:
CC    A;Gene: Anla
CC    C;Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
CC    C;Keywords: fusion protein; zinc; zinc finger
CC    F;28-103/Region: ubiquitin-like protein
CC    F;28-103/Domain: ubiquitin homology <UBH>
```

```
CC    F;625-693/Region: zinc finger
SQ   SEQUENCE   693 AA; 76844 MW; 2512772 CN;

  Query Match           65.9%;  Score 54;  DB 2;  Length 693;
  Best Local Similarity  50.0%;  Pred. No. 1.04e+02;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    474 LFRSVEVRNIAD 485
          |||:|  :::||
Qy      1 LFRAVITKKVAD 12


RESULT    5
ID   JN0674     STANDARD;     PRT;   701 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: jn0674  check: 3367  from: 1  to: 701.
XX
CC     TOIG of: jn0674  check: 3367  from: 1  to: 701
CC
CC    >P1;JN0674
CC    ubiquitin-like fusion protein An1b - African clawed frog
CC    C;Species: Xenopus laevis (African clawed frog)
CC    C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 08-Sep-19
  97
CC    C;Accession: JN0674
CC    R;Linnen, J.M.; Bailey, C.P.; Weeks, D.L.
CC    Gene 128, 181-188, 1993
CC    A;Title: Two related localized mRNAs from Xenopus laevis encode ubiquitin
  -like fusion proteins.
CC    A;Reference number: JN0673
CC    A;Accession: JN0674
CC    A;Status: nucleic acid sequence not shown
CC    A;Molecule type: mRNA
CC    A;Residues: 1-701 <LIN>
CC    C;Genetics:
CC    A;Gene: An1b
CC    C;Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
CC    C;Keywords: fusion protein; zinc; zinc finger
CC    F;28-103/Region: ubiquitin-like protein
CC    F;28-103/Domain: ubiquitin homology <UBH>
CC    F;633-701/Region: zinc finger
SQ   SEQUENCE   701 AA; 78581 MW; 2611877 CN;

  Query Match           65.9%;  Score 54;  DB 2;  Length 701;
  Best Local Similarity  50.0%;  Pred. No. 1.04e+02;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    481 LFRSVEVRNIAD 492
          |||:|  :::||
Qy      1 LFRAVITKKVAD 12


RESULT    6
ID   S59660     STANDARD;     PRT;   885 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s59660  check: 8781  from: 1  to: 885.
XX
CC     TOIG of: s59660  check: 8781  from: 1  to: 885
CC
CC    >P1;S59660
CC    anaphase spindle elongation protein ASE1 - yeast (Saccharomyces cerevisia
  e)
CC    N;Alternate names: protein O2806; protein YOR058c
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Dec-19
  97
CC    C;Accession: S59660; S66941
CC    R;Pellman, D.; Fink, G.R.
CC    submitted to the EMBL Data Library, January 1995
CC    A;Description: Yeast microtubule-associated proteins required for anaphas
  e spindle elongation.
CC    A;Reference number: S59660
CC    A;Accession: S59660
CC    A;Molecule type: DNA
CC    A;Residues: 1-885 <PEL>
CC    A;Cross-references: EMBL:U20235; NID:g972941; PID:g972942
CC    R;Bohn, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens
  , M.
CC    submitted to the Protein Sequence Database, July 1996
CC    A;Reference number: S66929
CC    A;Accession: S66941
CC    A;Molecule type: DNA
CC    A;Residues: 1-885 <BOH>
CC    A;Cross-references: EMBL:Z74966; NID:g1420196; PID:e252338; PID:g1420197;
   MIPS:YOR058c
CC    A;Experimental source: strain S288C
CC    C;Genetics:
CC    A;Gene: SGD:ASE1
CC    A;Cross-references: SGD:S0005584
CC    A;Map position: 15R
SQ   SEQUENCE   885 AA; 101623 MW; 3976412 CN;

  Query Match           65.9%;  Score 54;  DB 2;  Length 885;
  Best Local Similarity  60.0%;  Pred. No. 1.04e+02;
  Matches     6;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    328 FKSVLTKKVS 337
          |::|:||||:
Qy      2 FRAVITKKVA 11


RESULT    7
ID   F64964     STANDARD;     PRT;   236 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: The complete genome sequence of Escherichia coli K-12.
XX
CC    A;Title: The complete genome sequence of Escherichia coli K-12.
CC    A;Reference number: A64720; MUID:97426617
CC    A;Accession: F64964
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-236 <BLAT>
CC    A;Cross-references: GB:AE000291; GB:U00096; NID:g1788298; PID:g1788308; U
  WGP:b1999
CC    A;Experimental source: strain K-12, substrain MG1655
SQ   SEQUENCE   236 AA; 27074 MW; 286157 CN;

  Query Match           64.6%;  Score 53;  DB 2;  Length 236;
  Best Local Similarity  50.0%;  Pred. No. 1.33e+02;
  Matches     6;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db      4 LFRVVISEQIID 15
          ||| ||:  : |
Qy      1 LFRAVITKKVAD 12


RESULT    8
ID   G64840     STANDARD;     PRT;   904 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
```

```
DE   A;Title: The complete genome sequence of Escherichia coli K-12.
XX
CC    A;Title: The complete genome sequence of Escherichia coli K-12.
CC    A;Reference number: A64720; MUID:97426617
CC    A;Accession: G64840
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-904 <BLAT>
CC    A;Cross-references: GB:AE000201; GB:U00096; NID:g2367113; PID:g1787227; U
  WGP:b0993
CC    A;Experimental source: strain K-12, substrain MG1655
CC    C;Genetics:
CC    A;Gene: torS
CC    C;Keywords: phosphotransferase
SQ   SEQUENCE   904 AA; 99738 MW; 3808259 CN;

  Query Match           64.6%;  Score 53;  DB 2;  Length 904;
  Best Local Similarity  54.5%;  Pred. No. 1.33e+02;
  Matches    6;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    766 LFRGIIPKPVP 776
          |||::| | |:
Qy      1 LFRAVITKKVA 11


RESULT    9
ID   S64131     STANDARD;     PRT;   126 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s64131  check: 414  from: 1  to: 126.
XX
CC     TOIG of: s64131  check: 414  from: 1  to: 126
CC
CC    >P1;S64131
CC    hypothetical protein YGL121c - yeast (Saccharomyces cerevisiae)
CC    N;Alternate names: hypothetical protein G2913
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 14-Nov-19
  97
CC    C;Accession: S64131
CC    R;Lauquin, G.
CC    submitted to the Protein Sequence Database, May 1996
CC    A;Reference number: S64122
CC    A;Accession: S64131
CC    A;Molecule type: DNA
CC    A;Residues: 1-126 <LAU>
CC    A;Cross-references: EMBL:Z72643; NID:g1322678; PID:e243349; PID:g1322679;
   MIPS:YGL121c
CC    A;Experimental source: strain S288C
CC    C;Genetics:
CC    A;Map position: 7L
SQ   SEQUENCE   126 AA; 14922 MW; 82947 CN;

  Query Match           63.4%;  Score 52;  DB 2;  Length 126;
  Best Local Similarity  33.3%;  Pred. No. 1.70e+02;
  Matches    4;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     82 LYRDVLSMKMSE 93
          |:| |:::|:::
Qy      1 LFRAVITKKVAD 12


RESULT   10
ID   JC2359     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: PH1298.
XX
CC    A;Accession: PH1298
CC    A;Molecule type: DNA
CC    A;Residues: 169-177 <TRA>
CC    R;Fenton, R.G.
CC    submitted to the EMBL Data Library, June 1994
CC    A;Reference number: G07128
CC    A;Accession: G01446
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: mRNA
CC    A;Residues: 1-317 <FEN>
CC    A;Cross-references: EMBL:U10340; NID:g499123; PID:g499124
CC    C;Genetics:
CC    A;Gene: MAGE-X2
CC    F;169-177/Region: HLA-A1 binding #status predicted
SQ   SEQUENCE   317 AA; 34928 MW; 530585 CN;

  Query Match           63.4%;  Score 52;  DB 2;  Length 317;
  Best Local Similarity  41.7%;  Pred. No. 1.70e+02;
  Matches    5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    105 LFREALSNKVDE 116
          |||  :::|| :
Qy      1 LFRAVITKKVAD 12


RESULT   11
ID   I38661     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: PH1297.
XX
CC    A;Accession: PH1297
CC    A;Molecule type: DNA
CC    A;Residues: 169-177 <TRA>
CC    C;Genetics:
CC    A;Gene: GDB:MAGE4
CC    A;Cross-references: GDB:331119
CC    A;Map position: Xq28-Xq28
CC    A;Introns: #status absent
SQ   SEQUENCE   317 AA; 34899 MW; 528124 CN;

  Query Match           63.4%;  Score 52;  DB 2;  Length 317;
  Best Local Similarity  41.7%;  Pred. No. 1.70e+02;
  Matches    5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    105 LFREALSNKVDE 116
          |||  :::|| :
Qy      1 LFRAVITKKVAD 12


RESULT   12
ID   S40731     STANDARD;     PRT;   489 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s40731  check: 4937  from: 1  to: 489.
XX
CC     TOIG of: s40731  check: 4937  from: 1  to: 489
CC
CC    >P1;S40731
CC    ATP-dependent RNA helicase homolog T26G10.1 - Caenorhabditis elegans
CC    C;Species: Caenorhabditis elegans
CC    C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-19
  97
CC    C;Accession: S40731
CC    R;Berks, M.
```

```
CC    submitted to the EMBL Data Library, December 1993
CC    A;Reference number: S40731
CC    A;Accession: S40731
CC    A;Molecule type: DNA
CC    A;Residues: 1-489 <BER>
CC    A;Cross-references: EMBL:Z29115; NID:g439259; PID:g439260
CC    C;Genetics:
CC    A;Introns: 26/1; 319/3; 403/1
SQ   SEQUENCE   489 AA; 54227 MW; 1183591 CN;

  Query Match           63.4%;  Score 52;  DB 2;  Length 489;
  Best Local Similarity  63.6%;  Pred. No. 1.70e+02;
  Matches     7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    223 LFSATMTKKVS 233
          || | :||||:
Qy      1 LFRAVITKKVA 11


RESULT   13
ID   I38667     STANDARD;     PRT;   234 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: Structure, chromosomal localization, and expression of 12 genes o
  f the MAGE family.
XX
CC    A;Title: Structure, chromosomal localization, and expression of 12 genes
  of the MAGE family.
CC    A;Reference number: I38659; MUID:95012457
CC    A;Accession: I38667
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-234 <RES>
CC    A;Cross-references: EMBL:U10693; NID:g533525; PID:g533526
CC    C;Genetics:
CC    A;Gene: GDB:MAGE8
CC    A;Cross-references: GDB:331123
CC    A;Map position: Xq28-Xq28
CC    A;Introns: #status absent
SQ   SEQUENCE   234 AA; 25197 MW; 296950 CN;

  Query Match           62.2%;  Score 51;  DB 2;  Length 234;
  Best Local Similarity  50.0%;  Pred. No. 2.16e+02;
  Matches     6;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db    107 LFREALDEKVAE 118
          |||  :  |||:
Qy      1 LFRAVITKKVAD 12


RESULT   14
ID   A27477     STANDARD;     PRT;   736 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   C;Genetics:.
XX
CC    C;Genetics:
CC    A;Gene: SGD:CDC24; CLS4; TSL1
CC    A;Cross-references: SGD:S0000039
CC    A;Map position: 1L
CC    C;Superfamily: CDC24 homology
CC    F;160-336/Domain: CDC24 homology <CD24>
SQ   SEQUENCE   736 AA; 83960 MW; 2862170 CN;

  Query Match           62.2%;  Score 51;  DB 2;  Length 736;
  Best Local Similarity  63.6%;  Pred. No. 2.16e+02;
  Matches     7;  Conservative    1;  Mismatches   3;  Indels   0;  Gaps   0;

Db    399 LFSEVVTKKSA 409
          ||  |:||| |
Qy      1 LFRAVITKKVA 11


RESULT   15
ID   S45444     STANDARD;     PRT;   980 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s45444  check: 8660  from: 1  to: 980.
XX
CC     TOIG of: s45444  check: 8660  from: 1  to: 980
CC
CC    >P1;S45444
CC    BEM1 protein-binding protein BOB1 - yeast (Saccharomyces cerevisiae)
CC    N;Alternate names: protein YBL0717; protein YBL085w
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 09-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 05-Dec-19
  97
CC    C;Accession: S45444; S45421; S45826; S59218
CC    R;Bender, A.; Bender, L.; Kokojan, V.
CC    submitted to the EMBL Data Library, April 1994
CC    A;Description: Yeast Bob1p (Bem1p-binding protein) binds to the SH3 domai
  n-containing protein Bem1p.
CC    A;Reference number: S45444
CC    A;Accession: S45444
CC    A;Molecule type: DNA
CC    A;Residues: 1-980 <BEN>
CC    A;Cross-references: EMBL:L31406; NID:g829041; PID:g466436
CC    R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
CC    submitted to the EMBL Data Library, May 1994
CC    A;Description: Sequence analysis of a 78,6 kb segment of the left end of
  Saccaromyces cerevisiae chromosome II.
CC    A;Reference number: S45387
CC    A;Accession: S45421
CC    A;Molecule type: DNA
CC    A;Residues: 1-980 <OBE>
CC    A;Cross-references: EMBL:X79489; NID:g496661; PID:g496694
CC    R;Domdey, H.; Gassenhuber, H.; Obermaier, B.; Piravandi, E.
CC    submitted to the Protein Sequence Database, August 1994
CC    A;Reference number: S45816
CC    A;Accession: S45826
CC    A;Molecule type: DNA
CC    A;Residues: 1-980 <DOM>
CC    A;Cross-references: EMBL:Z35846; NID:g536137; PID:g536138; MIPS:YBL085w
CC    R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
CC    Yeast 11, 1103-1112, 1995
CC    A;Title: Sequence analysis of a 78.6 kb segment of the left end of Saccha
  romyces cerevisiae chromosome II.
CC    A;Reference number: S59184
CC    A;Accession: S59218
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-980 <OBW>
CC    A;Cross-references: EMBL:X79489; NID:g496661; PID:g496694
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, M
  ay 1994
CC    C;Genetics:
CC    A;Gene: SGD:BOI1; BOB1
CC    A;Cross-references: MIPS:YBL085w; SGD:S0000181
CC    A;Map position: 2L
CC    C;Superfamily: SH3 homology
CC    F;20-72/Domain: SH3 homology <SH3>
SQ   SEQUENCE   980 AA; 109294 MW; 5181201 CN;

  Query Match           62.2%;  Score 51;  DB 2;  Length 980;
  Best Local Similarity  54.5%;  Pred. No. 2.16e+02;
  Matches     6;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;
```

```
Db     66 LYPAVFTKRIA 76
          |: ||:||::|
Qy      1 LFRAVITKKVA 11


RESULT   16
ID   S57596     STANDARD;     PRT;   1729 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s57596  check: 3679  from: 1  to: 1729.
XX
CC     TOIG of: s57596  check: 3679  from: 1  to: 1729
CC
CC    >P1;S57596
CC    hypothetical protein YMR229c - yeast (Saccharomyces cerevisiae)
CC    N;Alternate names: hypothetical protein YM9959.11c
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Dec-19
  97
CC    C;Accession: S57596
CC    R;Skelton, J.; Churcher, C.M.
CC    submitted to the EMBL Data Library, June 1995
CC    A;Reference number: S57587
CC    A;Accession: S57596
CC    A;Molecule type: DNA
CC    A;Residues: 1-1729 <SKE>
CC    A;Cross-references: EMBL:Z49939; NID:g887599; PID:g887610; MIPS:YMR229c
CC    A;Experimental source: strain AB972
CC    C;Genetics:
CC    A;Gene: SGD:FMI1
CC    A;Cross-references: SGD:S0004842
CC    A;Map position: 13R
SQ   SEQUENCE   1729 AA; 193133 MW; 15516437 CN;

  Query Match           62.2%;  Score 51;  DB 2;  Length 1729;
  Best Local Similarity  54.5%;  Pred. No. 2.16e+02;
  Matches     6;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db   1674 LFERIITKKIT 1684
          ||  :||||::
Qy      1 LFRAVITKKVA 11


RESULT   17
ID   S18521     STANDARD;     PRT;   348 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S18521.
XX
CC    A;Accession: S18521
CC    A;Molecule type: DNA
CC    A;Residues: 1-348 <KIT>
CC    A;Cross-references: EMBL:X61672; NID:g4977; PID:g4978
CC    C;Keywords: G protein-coupled receptor; transmembrane protein
SQ   SEQUENCE   348 AA; 39285 MW; 700740 CN;

  Query Match           61.0%;  Score 50;  DB 2;  Length 348;
  Best Local Similarity  55.6%;  Pred. No. 2.75e+02;
  Matches     5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    229 LFRAILIRK 237
          ||||:: :|
Qy      1 LFRAVITKK 9


RESULT   18
ID   YQBZDH     STANDARD;     PRT;   394 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: yqbzdh  check: 2838  from: 1  to: 394.
XX
CC     TOIG of: yqbzdh  check: 2838  from: 1  to: 394
CC
CC    >P1;YQBZDH
CC    fimD protein - Dichelobacter nodosus (serotype H1)
CC    C;Species: Dichelobacter nodosus
CC    C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 05-Sep-19
  97
CC    C;Accession: S15255
CC    R;Hobbs, M.; Dalrymple, B.P.; Cox, P.T.; Livingstone, S.P.; Delaney, S.F.
  ; Mattick, J.S.
CC    Mol. Microbiol. 5, 543-560, 1991
CC    A;Title: Organization of the fimbrial gene region of Bacteroides nodosus:
   class I and class II strains.
CC    A;Reference number: S15240; MUID:91260439
CC    A;Accession: S15255
CC    A;Molecule type: DNA
CC    A;Residues: 1-394 <HOB>
CC    A;Cross-references: EMBL:X52390; NID:g39703; PID:g580812
CC    A;Note: the source is designated as Bacteroides nodosus
CC    C;Genetics:
CC    A;Gene: fimD
CC    A;Start codon: GTG
CC    C;Superfamily: fimD protein
CC    C;Keywords: fimbria
SQ   SEQUENCE   394 AA; 45105 MW; 815478 CN;

  Query Match           61.0%;  Score 50;  DB 1;  Length 394;
  Best Local Similarity  72.7%;  Pred. No. 2.75e+02;
  Matches     8;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db    365 FRAASTKKTAD 375
          |||  ||| ||
Qy      2 FRAVITKKVAD 12


RESULT   19
ID   S64646     STANDARD;     PRT;   690 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s64646  check: 478  from: 1  to: 690.
XX
CC     TOIG of: s64646  check: 478  from: 1  to: 690
CC
CC    >P1;S64646
CC    biotin--[acetyl-CoA-carboxylase] ligase (EC 6.3.4.15) - yeast (Saccharomy
  ces cerevisiae)
CC    N;Alternate names: biotin-apoprotein ligase; protein D2140; protein YDL14
  1w
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 12-Jul-1996 #sequence_revision 23-Aug-1996 #text_change 05-Dec-19
  97
CC    C;Accession: S64646; S67687; S67688
CC    R;Cronan Jr., J.E.; Wallace, J.C.
CC    FEMS Microbiol. Lett. 130, 221-230, 1995
CC    A;Title: The gene encoding the biotin-apoprotein ligase of Saccharomyces
  cerevisiae.
CC    A;Reference number: S64646
CC    A;Accession: S64646
CC    A;Molecule type: DNA
CC    A;Residues: 1-690 <CRO>
CC    A;Cross-references: EMBL:U27182; NID:g886080; PID:g886081
CC    R;Saluz, H.P.; Woelfl, S.; Hanemann, V.
CC    submitted to the Protein Sequence Database, July 1996
```

```
CC    A;Reference number: S67677
CC    A;Accession: S67687
CC    A;Molecule type: DNA
CC    A;Residues: 1-690 <SAL>
CC    A;Cross-references: EMBL:Z74189; NID:g1431218; PID:e253238; PID:g1431219;
   MIPS:YDL141w
CC    A;Experimental source: strain S288C
CC    R;Baron, L.; Legros, Y.; Biteau, N.; Monnet, A.; Granotier, C.
CC    submitted to the Protein Sequence Database, July 1996
CC    A;Reference number: S67688
CC    A;Accession: S67688
CC    A;Molecule type: DNA
CC    A;Residues: 1-291 <BAR>
CC    A;Cross-references: EMBL:Z74189; MIPS:YDL141w
CC    A;Experimental source: strain S288C
CC    C;Genetics:
CC    A;Gene: SGD:BPL1
CC    A;Cross-references: SGD:S0002300
CC    A;Map position: 4L
CC    C;Keywords: ligase; transmembrane protein
CC    F;458-474/Domain: transmembrane #status predicted <TMM>
SQ   SEQUENCE   690 AA; 76362 MW; 2564839 CN;

  Query Match           61.0%;  Score 50;  DB 2;  Length 690;
  Best Local Similarity  50.0%;  Pred. No. 2.75e+02;
  Matches     5;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db    679 IFKSLIAKKV 688
          :|:::|:|||
Qy      1 LFRAVITKKV 10


RESULT   20
ID   F64991     STANDARD;     PRT;   890 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: The complete genome sequence of Escherichia coli K-12.
XX
CC    A;Title: The complete genome sequence of Escherichia coli K-12.
CC    A;Reference number: A64720; MUID:97426617
CC    A;Accession: F64991
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-890 <BLAT>
CC    A;Cross-references: GB:AE000310; GB:U00096; NID:g2367131; PID:g1788545; U
  WGP:b2216
CC    A;Experimental source: strain K-12, substrain MG1655
CC    C;Genetics:
CC    A;Gene: yojN
SQ   SEQUENCE   890 AA; 100371 MW; 3869027 CN;

  Query Match           61.0%;  Score 50;  DB 2;  Length 890;
  Best Local Similarity  60.0%;  Pred. No. 2.75e+02;
  Matches     6;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    377 RTVISNKIAD 386
          |:||::|:||
Qy      3 RAVITKKVAD 12


RESULT   21
ID   I40457     STANDARD;     PRT;   2560 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: i40457  check: 9702  from: 1  to: 2560.
XX
CC    TOIG of: i40457  check: 9702  from: 1  to: 2560
CC
CC    >P1;I40457
CC    peptide-synthetase 2 - Bacillus subtilis
CC    N;Alternate names: pps2
CC    C;Species: Bacillus subtilis
CC    C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 18-Jul-19
  97
CC    C;Accession: I40457; S49134
CC    R;Tognoni, A.; Franchi, E.; Magistrelli, C.; Colombo, E.; Cosmina, P.; Gr
  andi, G.
CC    Microbiology 141, 645-648, 1995
CC    A;Title: A putative new peptide synthase operon in Bacillus subtilis: par
  tial characterization.
CC    A;Reference number: I40454; MUID:95227362
CC    A;Accession: I40457
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-2560 <RES>
CC    A;Cross-references: EMBL:Z34883; NID:g509465; PID:g509469
CC    C;Genetics:
CC    A;Gene: pps2
CC    C;Superfamily: gramicidin S synthetase I repeat homology; acetate--CoA li
  gase homology; acyl carrier protein homology
CC    F;456-1036/Domain: gramicidin S synthetase I repeat homology <GRS1>
CC    F;515-952/Domain: acetate--CoA ligase homology <ACL1>
CC    F;968-1036/Domain: acyl carrier protein homology <ACP1>
CC    F;1486-2076/Domain: gramicidin S synthetase I repeat homology <GRS2>
CC    F;1546-1992/Domain: acetate--CoA ligase homology <ACL2>
CC    F;2009-2076/Domain: acyl carrier protein homology <ACP2>
SQ   SEQUENCE   2560 AA; 290162 MW; 32837003 CN;

  Query Match           61.0%;  Score 50;  DB 2;  Length 2560;
  Best Local Similarity  41.7%;  Pred. No. 2.75e+02;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     63 IFRTIFIKEVPD 74
          :||::: | |:|
Qy      1 LFRAVITKKVAD 12


RESULT   22
ID   B69681     STANDARD;     PRT;   2560 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Status: preliminary; nucleic acid sequence not shown; translation not sh
  own.
XX
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-2560 <KUN>
CC    A;Experimental source: strain 168
CC    C;Genetics:
CC    A;Gene: ppsB
SQ   SEQUENCE   2560 AA; 290162 MW; 32837003 CN;

  Query Match           61.0%;  Score 50;  DB 2;  Length 2560;
  Best Local Similarity  41.7%;  Pred. No. 2.75e+02;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     63 IFRTIFIKEVPD 74
          :||::: | |:|
Qy      1 LFRAVITKKVAD 12


RESULT   23
ID   WMBEHG     STANDARD;     PRT;   224 AA.
XX
AC   xxxxxx
```

```
XX
DT   01-JAN-1900
XX
DE   A;Accession: JQ1494.
XX
CC    A;Accession: JQ1494
CC    A;Molecule type: DNA
CC    A;Residues: 1-224 <MCG>
CC    A;Cross-references: GB:D10470; DDBJ:D01127; NID:g221791; PID:d1001734; PI
  D:g221792
CC    C;Genetics:
CC    A;Gene: UL1
CC    C;Superfamily: varicella-zoster virus gene 60 protein
SQ   SEQUENCE   224 AA; 25192 MW; 242721 CN;

  Query Match           59.8%;  Score 49;  DB 1;  Length 224;
  Best Local Similarity  50.0%;  Pred. No. 3.49e+02;
  Matches    6;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db     27 VLRSVIAKEVGD 38
          ::|:||:| |:|
Qy      1 LFRAVITKKVAD 12


RESULT   24
ID   H64425     STANDARD;     PRT;   236 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Authors: Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
  C.R.; Venter, J.C.
XX
CC    A;Authors: Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
   C.R.; Venter, J.C.
CC    A;Title: Complete genome sequence of the methanogenic archaeon, Methanoco
  ccus jannaschii.
CC    A;Reference number: A64300; MUID:96337999
CC    A;Accession: H64425
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-236 <BUL>
CC    A;Cross-references: GB:U67543; GB:L77117; NID:g1591663; PID:g1591668; TIG
  R:MJ1009; PID:g1511039
CC    C;Genetics:
CC    A;Map position: REV939905-939195
SQ   SEQUENCE   236 AA; 27426 MW; 265381 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 236;
  Best Local Similarity  33.3%;  Pred. No. 3.49e+02;
  Matches    4;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    161 LYRSMISKIIEN 172
          |:|::|:| : :
Qy      1 LFRAVITKKVAD 12


RESULT   25
ID   S18857     STANDARD;     PRT;   274 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s18857  check: 6782  from: 1  to: 274.
XX
CC     TOIG of: s18857  check: 6782  from: 1  to: 274
CC
CC    >P1;S18857
CC    2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase (EC 2.3.1.
  117) - Actinobacillus pleuropneumoniae
CC    N;Alternate names: tetrahydrodipicolinate N-succinyltransferase
CC    C;Species: Actinobacillus pleuropneumoniae
CC    C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 31-Oct-19
  97
CC    C;Accession: S18857
CC    R;Denich, K.; O'Hanley, P.; Lalonde, G.
CC    submitted to the EMBL Data Library, November 1991
CC    A;Description: Cloning and sequence analysis of the DAPD gene from Actino
  bacillus pleuropneumonia.
CC    A;Reference number: S18857
CC    A;Accession: S18857
CC    A;Status: preliminary
CC    A;Molecule type: DNA
CC    A;Residues: 1-274 <DEN>
CC    A;Cross-references: EMBL:X63201; NID:g38946; PID:g38947
CC    C;Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransfe
  rase
CC    C;Keywords: acyltransferase; aminotransferase; diaminopimelate biosynthes
  is; lysine biosynthesis; pyridoxal phosphate
SQ   SEQUENCE   274 AA; 29761 MW; 412262 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 274;
  Best Local Similarity  70.0%;  Pred. No. 3.49e+02;
  Matches    7;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    246 LYCAVIVKKV 255
          |: ||| |||
Qy      1 LFRAVITKKV 10


RESULT   26
ID   XNECSD     STANDARD;     PRT;   274 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: The complete genome sequence of Escherichia coli K-12.
XX
CC    A;Title: The complete genome sequence of Escherichia coli K-12.
CC    A;Reference number: A64720; MUID:97426617
CC    A;Accession: F64740
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-274 <BLAT>
CC    A;Cross-references: GB:AE000126; GB:U00096; NID:g1786358; PID:g1786362; U
  WGP:b0166
CC    A;Experimental source: strain K-12, substrain MG1655
CC    R;Richaud, C.; Richaud, F.; Martin, C.; Haziza, C.; Patte, J.C.
CC    J. Biol. Chem. 259, 14824-14828, 1984
CC    A;Title: Regulation of expression and nucleotide sequence of the Escheric
  hia coli dapD gene.
CC    A;Reference number: A00601; MUID:85054973
CC    A;Accession: A00601
CC    A;Molecule type: DNA
CC    A;Residues: 1-30,'D',32-162,'R',164-176,'M',178-189,'L',191-274 <RIC>
CC    R;Fujita, N.
CC    submitted to the EMBL Data Library, January 1994
CC    A;Reference number: S45181
CC    A;Accession: S45231
CC    A;Status: preliminary
CC    A;Molecule type: DNA
CC    A;Residues: 1-274 <FUJ>
CC    A;Cross-references: EMBL:D26562; NID:g473770; PID:d1006154; PID:g473821
CC    R;van Heeswijk, W.C.; Rabenberg, M.; Westerhoff, H.V.; Kahn, D.
CC    Mol. Microbiol. 9, 443-457, 1993
CC    A;Title: The genes of the glutamine synthetase adenylylation cascade are
  not regulated by nitrogen in Escherichia coli.
CC    A;Reference number: S36254
CC    A;Accession: S36255
CC    A;Molecule type: DNA
CC    A;Residues: 1-14 <VAN>
```

```
CC    A;Cross-references: EMBL:Z21842
CC    C;Genetics:
CC    A;Gene: dapD
CC    A;Map position: 4 min
CC    C;Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransfe
  rase
CC    C;Keywords: acyltransferase; diaminopimelate biosynthesis; lysine biosynt
  hesis; pyridoxal phosphate
SQ   SEQUENCE   274 AA; 29892 MW; 403030 CN;

  Query Match           59.8%;  Score 49;  DB 1;  Length 274;
  Best Local Similarity  70.0%;  Pred. No. 3.49e+02;
  Matches     7;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    247 LYCAVIVKKV 256
          |: ||| |||
Qy      1 LFRAVITKKV 10


RESULT   27
ID   H64538     STANDARD;     PRT;   287 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Authors: Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
  C.M.; Venter, J.C.
XX
CC    A;Authors: Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
   C.M.; Venter, J.C.
CC    A;Title: The complete genome sequence of the gastric pathogen Helicobacte
  r pylori.
CC    A;Reference number: A64520; MUID:97394467
CC    A;Accession: H64538
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-287 <TOM>
CC    A;Cross-references: GB:AE000536; GB:AE000511; NID:g2313230; PID:g2313242;
   TIGR:HP0152
CC    C;Genetics:
CC    A;Start codon: TTG
SQ   SEQUENCE   287 AA; 32948 MW; 428117 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 287;
  Best Local Similarity  55.6%;  Pred. No. 3.49e+02;
  Matches     5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    189 LYRAILIKK 197
          |:||:: ||
Qy      1 LFRAVITKK 9


RESULT   28
ID   H64133     STANDARD;     PRT;   303 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Authors: Cotton, M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Saude
  k, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.
XX
CC    A;Authors: Cotton, M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Saud
  ek, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L
CC    Fuhrmann, J.L.; Geoghagen, N.S.M.; Gnehm, C.L.; McDonald, L.A.; Small, K.
  V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.
CC    A;Title: Whole-genome random sequencing and assembly of Haemophilus influ
  enzae Rd.
CC    A;Reference number: A64000; MUID:95350630
CC    A;Accession: H64133
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-303 <TIGR>
CC    A;Cross-references: GB:U32836; GB:L42023; NID:g1574473; PID:g1574480; TIG
  R:HI1634
CC    C;Genetics:
CC    A;Start codon: GTG
CC    C;Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransfe
  rase
CC    C;Keywords: acyltransferase; diaminopimelate biosynthesis; lysine biosynt
  hesis; pyridoxal phosphate
SQ   SEQUENCE   303 AA; 32817 MW; 503347 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 303;
  Best Local Similarity  70.0%;  Pred. No. 3.49e+02;
  Matches     7;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    275 LYCAVIVKKV 284
          |: ||| |||
Qy      1 LFRAVITKKV 10


RESULT   29
ID   MTTILTSTFRNLSTTSKWALRFSVRPLSCSSQVQSAPAVQTKSKKTLAKPNLKNIVVVEGVRIPFLLSGT
   STANDARD;     PRT;   405 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   This is a DE line.
XX
SQ   SEQUENCE   405 AA; 43796 MW; 794088 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 405;
  Best Local Similarity  50.0%;  Pred. No. 3.49e+02;
  Matches     6;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db     19 LYRTNIPKDVVD 30
          |:|: | | | |
Qy      1 LFRAVITKKVAD 12


RESULT   30
ID   A54926     STANDARD;     PRT;   468 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: a54926  check: 8551  from: 1  to: 468.
XX
CC     TOIG of: a54926  check: 8551  from: 1  to: 468
CC
CC    >P1;A54926
CC    UDPglucose 6-dehydrogenase (EC 1.1.1.22) - bovine
CC    C;Species: Bos primigenius taurus (cattle)
CC    C;Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 04-Nov-19
  94
CC    C;Accession: A54926
CC    R;Hempel, J.; Perozich, J.; Romovacek, H.; Hinich, A.; Kuo, I.; Feingold,
   D.S.
CC    Protein Sci. 3, 1074-1080, 1994
CC    A;Title: UDP-glucose dehydrogenase from bovine liver: primary structure a
  nd relationship to other dehydrogenases.
CC    A;Reference number: A54926
CC    A;Accession: A54926
CC    A;Status: preliminary
CC    A;Molecule type: protein
CC    A;Residues: 1-468 <HEM>
CC    C;Keywords: NAD; oxidoreductase
SQ   SEQUENCE   468 AA; 52171 MW; 1108556 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 468;
```

```
  Best Local Similarity  54.5%;  Pred. No. 3.49e+02;
  Matches     6;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db    321 LFNTVTDKKIA 331
          || :|  ||:|
Qy      1 LFRAVITKKVA 11


RESULT   31
ID   S32593     STANDARD;     PRT;   528 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s32593  check: 6953  from: 1  to: 528.
XX
CC     TOIG of: s32593  check: 6953  from: 1  to: 528
CC
CC    >P1;S32593
CC    [RNA-polymerase]-subunit kinase (EC 2.7.1.141) alpha chain - yeast (Sacch
  aromyces cerevisiae)
CC    N;Alternate names: CTD kinase alpha chain; CTD kinase largest chain; prot
  ein YKL139w
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-Dec-19
  97
CC    C;Accession: S32593; S37968; S27423
CC    R;Lee, J.M.; Greenleaf, A.L.
CC    Gene Expr. 1, 149-167, 1991
CC    A;Title: CTD kinase large subunit is encoded by CTK1, a gene required for
   normal growth of Saccharomyces cerevisiae.
CC    A;Reference number: S32593
CC    A;Accession: S32593
CC    A;Molecule type: DNA
CC    A;Residues: 1-528 <LEE>
CC    A;Cross-references: EMBL:M69024; NID:g171327; PID:g171328
CC    R;Ramezani Rad, M.; Xu, G.; Kirchrath, L.; Fritz, C.; Keuchel, H.; Hollen
  berg, C.P.
CC    submitted to the Protein Sequence Database, March 1994
CC    A;Reference number: S37953
CC    A;Accession: S37968
CC    A;Molecule type: DNA
CC    A;Residues: 1-528 <RAM>
CC    A;Cross-references: EMBL:Z28139; NID:g486234; PID:g486235; MIPS:YKL139w
CC    C;Genetics:
CC    A;Gene: SGD:CTK1
CC    A;Cross-references: SGD:S0001622
CC    A;Map position: 11L
CC    C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protei
  n kinase homology
CC    C;Keywords: ATP; autophosphorylation; nucleus; phosphoprotein; phosphotra
  nsferase; serine/threonine-specific protein kinase
CC    F;181-436/Domain: protein kinase homology <KIN>
CC    F;189-197/Region: protein kinase ATP-binding motif
CC    F;306/Active site: Asp #status predicted
SQ   SEQUENCE   528 AA; 60501 MW; 1441629 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 528;
  Best Local Similarity  66.7%;  Pred. No. 3.49e+02;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db    162 RSVIAKKVP 170
          |:||:|||:
Qy      3 RAVITKKVA 11


RESULT   32
ID   S40098     STANDARD;     PRT;   660 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s40098  check: 8490  from: 1  to: 660.
XX
CC     TOIG of: s40098  check: 8490  from: 1  to: 660
CC
CC    >P1;S40098
CC    hypothetical protein - Clostridium butyricum
CC    C;Species: Clostridium butyricum
CC    C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-19
  97
CC    C;Accession: S40098
CC    R;Hesslinger, C.; Muller, N.; Kaiser, M.; Sawers, G.
CC    submitted to the EMBL Data Library, December 1993
CC    A;Description: Identification of a gene encoding a second pyruvate format
  e-lyase in E. coli.
CC    A;Reference number: S40097
CC    A;Accession: S40098
CC    A;Status: preliminary
CC    A;Molecule type: DNA
CC    A;Residues: 1-660 <HES>
CC    A;Cross-references: EMBL:Z29084; NID:g436131; PID:g436133
SQ   SEQUENCE   660 AA; 78996 MW; 2241154 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 660;
  Best Local Similarity  62.5%;  Pred. No. 3.49e+02;
  Matches     5;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db    210 FRVIVTKK 217
          || ::|||
Qy      2 FRAVITKK 9


RESULT   33
ID   S53398     STANDARD;     PRT;   1454 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s53398  check: 1250  from: 1  to: 1454.
XX
CC     TOIG of: s53398  check: 1250  from: 1  to: 1454
CC
CC    >P1;S53398
CC    hypothetical protein YLR320w - yeast (Saccharomyces cerevisiae)
CC    N;Alternate names: hypothetical protein L8543.13
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 05-May-1995 #sequence_revision 01-Sep-1995 #text_change 21-Nov-19
  97
CC    C;Accession: S53398
CC    R;Du, Z.
CC    submitted to the EMBL Data Library, February 1995
CC    A;Description: The sequence of S. cerevisiae cosmid 8543.
CC    A;Reference number: S53390
CC    A;Accession: S53398
CC    A;Molecule type: DNA
CC    A;Residues: 1-1454 <DUZ>
CC    A;Cross-references: EMBL:U20618; NID:g2258165; PID:g662137; MIPS:YLR320w
CC    A;Experimental source: strain S288C (AB972)
CC    C;Genetics:
CC    A;Map position: 12R
SQ   SEQUENCE   1454 AA; 167681 MW; 10949283 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 1454;
  Best Local Similarity  41.7%;  Pred. No. 3.49e+02;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    623 IFRVIIDKGITD 634
          :|| :| | ::|
Qy      1 LFRAVITKKVAD 12


RESULT   34
```

```
ID   S68235     STANDARD;     PRT;   1906 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   F;1451-1708/Domain: protein kinase homology <KIN>.
XX
CC    F;1451-1708/Domain: protein kinase homology <KIN>
CC    F;1459-1467/Region: protein kinase ATP-binding motif
CC    F;1750-1906/Product: telokin #status predicted <TKN>
SQ   SEQUENCE   1906 AA; 210445 MW; 18691293 CN;

  Query Match           59.8%;  Score 49;  DB 2;  Length 1906;
  Best Local Similarity  50.0%;  Pred. No. 3.49e+02;
  Matches     5;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    909 FRDILGKKVS 918
          || :: |||:
Qy      2 FRAVITKKVA 11


RESULT   35
ID   C64328     STANDARD;     PRT;   193 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Authors: Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
  C.R.; Venter, J.C.
XX
CC    A;Authors: Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
   C.R.; Venter, J.C.
CC    A;Title: Complete genome sequence of the methanogenic archaeon, Methanoco
  ccus jannaschii.
CC    A;Reference number: A64300; MUID:96337999
CC    A;Accession: C64328
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-193 <BUL>
CC    A;Cross-references: GB:U67478; GB:L77117; NID:g1590958; PID:g1590963; TIG
  R:MJ0226; PID:g1510334
CC    C;Genetics:
CC    A;Map position: FOR216394-216975
CC    A;Start codon: TTG
SQ   SEQUENCE   193 AA; 22202 MW; 192384 CN;

  Query Match           58.5%;  Score 48;  DB 2;  Length 193;
  Best Local Similarity  25.0%;  Pred. No. 4.42e+02;
  Matches     3;  Conservative    7;  Mismatches   2;  Indels   0;  Gaps   0;

Db    129 LFKGIVKGRVSE 140
          ||:::: :|::
Qy      1 LFRAVITKKVAD 12


Search completed: Tue Apr  7 08:40:46 1998
Job time : 14 secs.
```

****

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:39:32 1998; MasPar time 3.95 Seconds
76.145 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-3
Description: (1-12) from 5541104.pep
Perfect Score: 82
Sequence: 1 LFRAVITKKVAD 12

Scoring table: PAM 150
Gap 15

Searched: 69112 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 25.999; Variance 27.975; scale 0.929

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	82	100.0	309	1	MAG1_HUMAN	MELANOMA-ASSOCIATED AN	7.13e-08
2	60	73.2	124	1	MAG5_HUMAN	MELANOMA-ASSOCIATED AN	4.58e-02
3	54	65.9	885	1	ASE1_YEAST	ANAPHASE SPINDLE ELONG	1.14e+00
4	53	64.6	236	1	YEEP_ECOLI	HYPOTHETICAL 27.1 KD P	1.89e+00
5	53	64.6	904	1	TORS_ECOLI	SENSOR PROTEIN TORS (E	1.89e+00
6	52	63.4	126	1	YGM1_YEAST	HYPOTHETICAL 14.9 KD P	3.13e+00
7	52	63.4	317	1	MAG4_HUMAN	MELANOMA-ASSOCIATED AN	3.13e+00
8	52	63.4	489	1	YN21_CAEEL	PUTATIVE ATP-DEPENDENT	3.13e+00
9	51	62.2	234	1	MAG8_HUMAN	MELANOMA-ASSOCIATED AN	5.14e+00
10	51	62.2	854	1	CC24_YEAST	CELL DIVISION CONTROL	5.14e+00
11	51	62.2	980	1	BOB1_YEAST	BOB1 PROTEIN (BEM1-BIN	5.14e+00
12	51	62.2	1729	1	RRP5_YEAST	RRNA BIOGENESIS PROTEI	5.14e+00
13	50	61.0	348	1	MAM2_SCHPO	PHEROMONE P-FACTOR REC	8.37e+00
14	50	61.0	394	1	FMDH_BACNO	POSSIBLE FIMBRIAL ASSE	8.37e+00
15	50	61.0	690	1	BPL1_YEAST	BIOTIN--PROTEIN LIGASE	8.37e+00
16	50	61.0	890	1	YOJN_ECOLI	PROBABLE SENSOR PROTEI	8.37e+00
17	50	61.0	2560	1	PPS2_BACSU	PEPTIDE SYNTHETASE 2.	8.37e+00
18	49	59.8	224	1	VGLL_HSV2H	GLYCOPROTEIN L PRECURS	1.35e+01
19	49	59.8	274	1	DAPD_ACTPL	2,3,4,5-TETRAHYDROPYRI	1.35e+01
20	49	59.8	274	1	DAPD_ECOLI	2,3,4,5-TETRAHYDROPYRI	1.35e+01
21	49	59.8	275	1	DAPD_HAEIN	2,3,4,5-TETRAHYDROPYRI	1.35e+01
22	49	59.8	432	1	GBA1_CRYNE	GUANINE NUCLEOTIDE-BIN	1.35e+01
23	49	59.8	475	1	ECHB_RAT	MITOCHONDRIAL TRIFUNCT	1.35e+01
24	49	59.8	528	1	CTK1_YEAST	CTD KINASE ALPHA SUBUN	1.35e+01
25	49	59.8	1906	1	KML2_CHICK	MYOSIN LIGHT CHAIN KIN	1.35e+01
26	48	58.5	193	1	Y226_METJA	HYPOTHETICAL PROTEIN M	2.16e+01
27	48	58.5	266	1	FLIP_CAUCR	FLAGELLAR BIOSYNTHETIC	2.16e+01
28	48	58.5	320	1	THI4_FUSOX	THIAZOLE BIOSYNTHETIC	2.16e+01
29	48	58.5	351	1	B3AR_CAVPO	BETA-3 ADRENERGIC RECE	2.16e+01
30	48	58.5	424	1	FADH_METMR	GLUTATHIONE-DEPENDENT	2.16e+01
31	48	58.5	516	1	RSP3_CHLRE	RADIAL SPOKE PROTEIN 3	2.16e+01
32	48	58.5	578	1	AMYM_BACLI	MALTOGENIC ALPHA-AMYLA	2.16e+01
33	48	58.5	754	1	YAJ3_SCHPO	PUTATIVE ATP-DEPENDENT	2.16e+01
34	47	57.3	69	1	RM39_YEAST	MITOCHONDRIAL 60S RIBO	3.43e+01
35	47	57.3	149	1	RL9_HAEIN	50S RIBOSOMAL PROTEIN	3.43e+01
36	47	57.3	173	1	CRAA_CERSI	ALPHA CRYSTALLIN A CHA	3.43e+01
37	47	57.3	173	1	CRAA_MUSVI	ALPHA CRYSTALLIN A CHA	3.43e+01
38	47	57.3	173	1	CRAA_EULFU	ALPHA CRYSTALLIN A CHA	3.43e+01
39	47	57.3	173	1	CRAA_PIG	ALPHA CRYSTALLIN A CHA	3.43e+01
40	47	57.3	173	1	CRAA_PERPO	ALPHA CRYSTALLIN A CHA	3.43e+01
41	47	57.3	225	1	UL92_EBV	PROTEIN BDLF4.	3.43e+01
42	47	57.3	249	1	H1_DROHY	HISTONE H1.	3.43e+01
43	47	57.3	301	1	THI2_SCHPO	THIAZOLE BIOSYNTHETIC	3.43e+01
44	47	57.3	314	1	MAG3_HUMAN	MELANOMA-ASSOCIATED AN	3.43e+01
45	47	57.3	353	1	GBQ_LYMST	GUANINE NUCLEOTIDE-BIN	3.43e+01
46	47	57.3	353	1	GBQ_CANFA	GUANINE NUCLEOTIDE-BIN	3.43e+01
47	47	57.3	353	1	GBQ_HUMAN	GUANINE NUCLEOTIDE-BIN	3.43e+01
48	47	57.3	353	1	GBQ3_DROME	GUANINE NUCLEOTIDE-BIN	3.43e+01
49	47	57.3	353	1	GBQ1_DROME	GUANINE NUCLEOTIDE-BIN	3.43e+01
50	47	57.3	353	1	GBQ_MOUSE	GUANINE NUCLEOTIDE-BIN	3.43e+01
51	47	57.3	353	1	GBQ_XENLA	GUANINE NUCLEOTIDE-BIN	3.43e+01
52	47	57.3	354	1	GBQ_LOLFO	GUANINE NUCLEOTIDE-BIN	3.43e+01
53	47	57.3	355	1	GB14_MOUSE	GUANINE NUCLEOTIDE-BIN	3.43e+01
54	47	57.3	359	1	GB11_XENLA	GUANINE NUCLEOTIDE-BIN	3.43e+01
55	47	57.3	359	1	GB11_HUMAN	GUANINE NUCLEOTIDE-BIN	3.43e+01
56	47	57.3	359	1	GB11_MOUSE	GUANINE NUCLEOTIDE-BIN	3.43e+01
57	47	57.3	359	1	GB11_MELGA	GUANINE NUCLEOTIDE-BIN	3.43e+01
58	47	57.3	370	1	FLGI_CAUCR	FLAGELLAR P-RING PROTE	3.43e+01
59	47	57.3	401	1	ALKB_PSEOL	ALKANE-1 MONOOXYGENASE	3.43e+01
60	47	57.3	441	1	FUS6_ARATH	FUSCA PROTEIN FUS6.	3.43e+01
61	47	57.3	468	1	ASPA_HELPY	ASPARTATE AMMONIA-LYAS	3.43e+01
62	47	57.3	523	1	TYD5_PAPSO	TYROSINE/DOPA DECARBOX	3.43e+01
63	47	57.3	648	1	Y084_HUMAN	HYPOTHETICAL PROTEIN K	3.43e+01
64	47	57.3	708	1	MR11_HUMAN	MRE11 HOMOLOG.	3.43e+01
65	47	57.3	774	1	KEMK_MOUSE	PUTATIVE SERINE/THREON	3.43e+01
66	47	57.3	1450	1	RPO1_ASFB7	DNA-DIRECTED RNA POLYM	3.43e+01
67	46	56.1	146	1	2SS_BEREX	2S SULPHUR-RICH SEED S	5.39e+01
68	46	56.1	206	1	RALB_HUMAN	RAS-RELATED PROTEIN RA	5.39e+01
69	46	56.1	215	1	CAT3_STAAU	CHLORAMPHENICOL ACETYL	5.39e+01
70	46	56.1	215	1	CAT2_STAAU	CHLORAMPHENICOL ACETYL	5.39e+01
71	46	56.1	215	1	CAT_STAIN	CHLORAMPHENICOL ACETYL	5.39e+01
72	46	56.1	215	1	CAT_STRAG	CHLORAMPHENICOL ACETYL	5.39e+01
73	46	56.1	274	1	T2B1_HERAU	TYPE II RESTRICTION EN	5.39e+01
74	46	56.1	285	1	WC2B_ARATH	PLASMA MEMBRANE INTRIN	5.39e+01
75	46	56.1	285	1	WC2C_ARATH	PLASMA MEMBRANE INTRIN	5.39e+01
76	46	56.1	324	1	THI4_FUSSH	THIAZOLE BIOSYNTHETIC	5.39e+01
77	46	56.1	327	1	GSPK_ECOLI	PUTATIVE GENERAL SECRE	5.39e+01
78	46	56.1	331	1	YZ44_METJA	HYPOTHETICAL PROTEIN M	5.39e+01
79	46	56.1	357	1	GBA2_DICDI	GUANINE NUCLEOTIDE-BIN	5.39e+01
80	46	56.1	369	1	KG1Z_YEAST	PROBABLE SERINE/THREON	5.39e+01
81	46	56.1	395	1	YKC8_CAEEL	PROBABLE NUCLEAR HORMO	5.39e+01
82	46	56.1	396	1	ASSY_METBA	ARGININOSUCCINATE SYNT	5.39e+01
83	46	56.1	432	1	AMIB_HAEIN	PROBABLE N-ACETYLMURAM	5.39e+01
84	46	56.1	489	1	CP56_YEAST	CYTOCHROME P450-DIT2 (	5.39e+01
85	46	56.1	548	1	ERF_HUMAN	ETS-RELATED PROTEIN ER	5.39e+01
86	46	56.1	564	1	BMRP_CANAL	BENOMYL/METHOTREXATE R	5.39e+01
87	46	56.1	593	1	PUR9_CHICK	PHOSPHORIBOSYLAMINOIMI	5.39e+01
88	46	56.1	697	1	ULA4_HCMVA	VIRION PROTEIN UL104.	5.39e+01
89	46	56.1	705	1	STT3_MOUSE	OLIGOSACCHARYL TRANSFE	5.39e+01
90	46	56.1	705	1	STT3_HUMAN	OLIGOSACCHARYL TRANSFE	5.39e+01
91	46	56.1	754	1	Y4QF_RHISN	PROBABLE PEPTIDASE Y4Q	5.39e+01
92	46	56.1	793	1	DCMA_METSO	CARBON MONOXIDE DEHYDR	5.39e+01
93	46	56.1	842	1	PHSH_VICFA	ALPHA-GLUCAN PHOSPHORY	5.39e+01
94	46	56.1	856	1	AAP1_YEAST	ALANINE/ARGININE AMINO	5.39e+01
95	46	56.1	863	1	YK44_YEAST	PUTATIVE 101.8 KD TRAN	5.39e+01
96	46	56.1	893	1	NIA_LEPMC	NITRATE REDUCTASE (NAD	5.39e+01

97	46	56.1	909	1	CT1A_FUSSO	CUTINASE TRANSCRIPTION	5.39e+01
98	46	56.1	1480	1	PAN1_YEAST	PAN1 PROTEIN.	5.39e+01
99	46	56.1	1556	1	GLTS_SYNY3	FERREDOXIN-DEPENDENT G	5.39e+01
100	45	54.9	235	1	COLI_RAT	CORTICOTROPIN-LIPOTROP	8.40e+01

ALIGNMENTS

```
RESULT    1
ID   MAG1_HUMAN     STANDARD;      PRT;   309 AA.
AC   P43355; O00346;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 1 (MAGE-1 ANTIGEN) (ANTIGEN MZ2-E).
GN   MAGEA1 OR MAGE1 OR MAGE1A.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92086861.
RA   VAN DER BRUGGEN P., TRAVERSARI C., CHOMEZ P., LURQUIN C., DE PLAEN E.,
RA   VAN DEN EYNDE B., KNUTH A., BOON T.;
RL   SCIENCE 254:1643-1647(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKIN;
RX   MEDLINE; 94311935.
RA   DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   GLOECKNER G., RUMP A., NORDSIEK G., HINZMANN B., KIOSCHIS P.,
RA   HEISS N., POUSTKA A., BAUER D., DRESCHER B., KNOB A., ROSENTHAL A.;
RL   SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [4]
RP   MUTAGENESIS.
RC   TISSUE=BLOOD;
RX   MEDLINE; 94157413.
RA   GAUGLER B., VAN DEN EYNDE B., VAN DER BRUGGEN P., ROMERO P.,
RA   GAFORIO J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;
RL   J. EXP. MED. 179:921-930(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC       CYTOLYTIC T LYMPHOCYTES.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES. NEVER EXPRESSED IN KIDNEY TUMORS, LEUKEMIAS AND
CC       LYMPHOMAS.
CC   -!- POLYMORPHISM: THE VARIANT AT POSITION 32 LIKELY REPRESENTS A
CC       POLYMORPHISM OF THE MAG-1 GENE.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; M77481; G416115; -.
DR   EMBL; U82672; G2078527; -.
DR   MIM; 300016; -.
KW   ANTIGEN; MULTIGENE FAMILY; POLYMORPHISM; TUMOR ANTIGEN.
FT   VARIANT      32     32       T -> A.
FT   DOMAIN       33     36       POLY-SER.
FT   MUTAGEN     163    163       D->A: ABOLISHES HLA-A1 BINDING.
FT   MUTAGEN     169    169       Y->A: ABOLISHES HLA-A1 BINDING.
FT   CONFLICT     72     72       R -> Q (IN REF. 3).
SQ   SEQUENCE   309 AA;  34342 MW;  E6CB1300 CRC32;

  Query Match          100.0%;  Score 82;  DB 1;  Length 309;
  Best Local Similarity 100.0%;  Pred. No. 7.13e-08;
  Matches    12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     97 LFRAVITKKVAD 108
         ||||||||||||
Qy      1 LFRAVITKKVAD 12
```

```
RESULT    2
ID   MAG5_HUMAN     STANDARD;      PRT;   124 AA.
AC   P43359;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 5 (MAGE-5 ANTIGEN).
GN   MAGEA5 OR MAGE5.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE TUMOR TRANSFORMATION
CC       OR PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; U10690; G533521; -.
DR   EMBL; U10689; G533519; -.
KW   ANTIGEN; MULTIGENE FAMILY.
FT   DOMAIN       40     43       POLY-SER.
SQ   SEQUENCE   124 AA;  13015 MW;  7216B8C8 CRC32;

  Query Match          73.2%;  Score 60;  DB 1;  Length 124;
  Best Local Similarity 66.7%;  Pred. No. 4.58e-02;
  Matches     8;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    104 VFRAALSKKVAD 115
          :||| ::|||||
Qy      1 LFRAVITKKVAD 12
```

```
RESULT    3
ID   ASE1_YEAST     STANDARD;      PRT;   885 AA.
AC   P50275;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   ANAPHASE SPINDLE ELONGATION PROTEIN.
GN   ASE1 OR YOR058C OR YOR29-09.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   PELLMAN D., FINK G.R.;
RL   SUBMITTED (SEP-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97279235.
RA   VALENS M., BOHN C., DAIGNAN-FORNIER B., DANG V., BOLOTIN-FUKUHARA M.;
RL   YEAST 13:379-390(1997).
CC   -!- FUNCTION: REQUIRED FOR ANAPHASE SPINDLE ELONGATION.
DR   EMBL; U20235; G972942; -.
DR   EMBL; Z74966; E252338; -.
DR   EMBL; Z70678; E234104; -.
DR   SGD; L0000125; ASE1.
SQ   SEQUENCE   885 AA;  101623 MW;  FF00B6B9 CRC32;

  Query Match          65.9%;  Score 54;  DB 1;  Length 885;
  Best Local Similarity 60.0%;  Pred. No. 1.14e+00;
  Matches     6;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    328 FKSVLTKKVS 337
```

```
          |::|:||||:
Qy      2 FRAVITKKVA 11


RESULT    4
ID   YEEP_ECOLI     STANDARD;      PRT;   236 AA.
AC   P76359; P94754;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 27.1 KD PROTEIN IN COBU-SBMC INTERGENIC REGION.
GN   YEEP.
OS   ESCHERICHIA COLI.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RA   BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL   SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE; 97251358.
RA   ITOH T., AIBA H., BABA T., FUJITA K., HAYASHI K., INADA T.,
RA   ISONO K., KASAI H., KIMURA S., KITAKAWA M., KITAGAWA M.,
RA   MAKINO K., MIKI T., MIZOBUCHI K., MORI H., MORI T., MOTOMURA K.,
RA   NAKADE S., NAKAMURA Y., NASHIMOTO H., NISHIO Y., OSHIMA T.,
RA   SAITO N., SAMPEI G., SEKI Y., SIVASUNDARAM S., TAGAMI H.,
RA   TAKEDA J., TAKEMOTO K., WADA C., YAMAMOTO Y., HORIUCHI T.;
RL   DNA RES. 3:379-392(1996).
CC   -!- SIMILARITY: TO E.COLI YFJP AND YKFA.
DR   EMBL; AE000291; G1788308; -.
DR   EMBL; D90838; G1736669; ALT_INIT.
DR   ECOGENE; EG13384; YEEP.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   236 AA;  27074 MW;  69582534 CRC32;

  Query Match           64.6%;  Score 53;  DB 1;  Length 236;
  Best Local Similarity  50.0%;  Pred. No. 1.89e+00;
  Matches     6;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db      4 LFRVVISEQIID 15
          ||| ||:  : |
Qy      1 LFRAVITKKVAD 12


RESULT    5
ID   TORS_ECOLI     STANDARD;      PRT;   904 AA.
AC   P39453; P75887;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   SENSOR PROTEIN TORS (EC 2.7.3.-).
GN   TORS.
OS   ESCHERICHIA COLI.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   MEJEAN V.;
RL   SUBMITTED (DEC-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RA   BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL   SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE OF 1-104 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE; 94364937.
RA   SIMON G., MEJEAN V., JOURLIN C., CHIPPAUX M., PASCAL M.-C.;
RL   J. BACTERIOL. 176:5601-5606(1994).
CC   -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM TORS/TORR.
CC       PROBABLY ACTIVATES TORR BY PHOSPHORYLATION.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC       (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC       KINASES.
DR   EMBL; X94231; E214041; -.
DR   EMBL; AE000201; G1787227; -.
DR   ECOGENE; EG12617; TORS.
KW   SENSORY TRANSDUCTION; TRANSFERASE; KINASE; PHOSPHORYLATION;
KW   TRANSMEMBRANE; INNER MEMBRANE.
FT   TRANSMEM    323    343       POTENTIAL.
FT   MOD_RES     443    443       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CONFLICT    761    762       QR -> HG (IN REF. 1).
SQ   SEQUENCE   904 AA;  99738 MW;  6F733AB3 CRC32;

  Query Match           64.6%;  Score 53;  DB 1;  Length 904;
  Best Local Similarity  54.5%;  Pred. No. 1.89e+00;
  Matches     6;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    766 LFRGIIPKPVP 776
          |||::| | |:
Qy      1 LFRAVITKKVA 11


RESULT    6
ID   YGM1_YEAST     STANDARD;      PRT;   126 AA.
AC   P53130;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 14.9 KD PROTEIN IN NAB2-ABC1 INTERGENIC REGION.
GN   YGL121C.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   LAUQUIN G.;
RL   SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z72643; E243349; -.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   126 AA;  14922 MW;  61575BF2 CRC32;

  Query Match           63.4%;  Score 52;  DB 1;  Length 126;
  Best Local Similarity  33.3%;  Pred. No. 3.13e+00;
  Matches     4;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db     82 LYRDVLSMKMSE 93
          |:| |:::|:::
Qy      1 LFRAVITKKVAD 12


RESULT    7
ID   MAG4_HUMAN     STANDARD;      PRT;   317 AA.
AC   P43358;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 4 (MAGE-4 ANTIGEN) (MAGE-X2) (MAGE-41).
GN   MAGEA4 OR MAGE4.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=SKIN;
RX   MEDLINE; 94311935.
RA   DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95369706.
RA   IMAI Y., SHICHIJO S., YAMADA A., KATAYAMA T., YANO H., ITOH K.;
RL   GENE 160:287-290(1995).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY WITH
CC       MAGE-1.
DR   EMBL; U10687; G533515; -.
DR   EMBL; U10688; G533517; -.
DR   EMBL; U10340; G499124; -.
DR   EMBL; D32077; G1125018; -.
KW   ANTIGEN; MULTIGENE FAMILY; POLYMORPHISM; TUMOR ANTIGEN.
FT   DOMAIN       41     44       POLY-SER.
FT   VARIANT     173    173       T -> A.
FT   CONFLICT    307    307       E -> Q (IN REF. 2).
SQ   SEQUENCE   317 AA;  34929 MW;  3CE38AF9 CRC32;

  Query Match           63.4%;  Score 52;  DB 1;  Length 317;
  Best Local Similarity  41.7%;  Pred. No. 3.13e+00;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    105 LFREALSNKVDE 116
          |||  :::|| :
Qy      1 LFRAVITKKVAD 12


RESULT    8
ID   YN21_CAEEL     STANDARD;      PRT;   489 AA.
AC   P34580;
DT   01-FEB-1994 (REL. 28, CREATED)
DT   01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT   01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE   PUTATIVE ATP-DEPENDENT RNA HELICASE T26G10.1 IN CHROMOSOME III.
GN   T26G10.1.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FRASER A.,
RA   FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
RA   JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISSTER N.,
RA   LATREILLE P., LIGHTNING J., LLOYD C., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SIMS M., SMALDON N., SMITH A., SMITH M., SONNHAMMER E., STADEN R.,
RA   SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K.,
RA   WATERSON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J.,
RA   WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
CC   -!- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.
CC   -!- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES.
DR   EMBL; Z29115; G439260; -.
DR   PIR; S40731; S40731.
DR   WORMPEP; T26G10.1; CE00337.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   HYPOTHETICAL PROTEIN; HELICASE; ATP-BINDING; RNA-BINDING.
FT   DOMAIN       31     40       ASP/GLU-RICH (ACIDIC).
FT-  NP_BIND      88     95       ATP (BY SIMILARITY).
FT   SITE *      194    197       DEAD BOX.
FT   DOMAIN      471    482       GLY-RICH.
SQ   SEQUENCE   489 AA;  54227 MW;  B9EFF81A CRC32;

  Query Match           63.4%;  Score 52;  DB 1;  Length 489;
  Best Local Similarity  63.6%;  Pred. No. 3.13e+00;
  Matches     7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    223 LFSATMTKKVS 233
          || | :||||:
Qy      1 LFRAVITKKVA 11


RESULT    9
ID   MAG8_HUMAN     STANDARD;      PRT;   234 AA.
AC   P43361;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 8 (MAGE-8 ANTIGEN).
GN   MAGEA8 OR MAGE8.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; U10693; G533526; -.
KW   ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT   DOMAIN       40     43       POLY-SER.
SQ   SEQUENCE   234 AA;  25197 MW;  D4931BC3 CRC32;

  Query Match           62.2%;  Score 51;  DB 1;  Length 234;
  Best Local Similarity  50.0%;  Pred. No. 5.14e+00;
  Matches     6;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db    107 LFREALDEKVAE 118
          |||  :  |||:
Qy      1 LFRAVITKKVAD 12


RESULT   10
ID   CC24_YEAST     STANDARD;      PRT;   854 AA.
AC   P11433;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   CELL DIVISION CONTROL PROTEIN 24 (CALCIUM REGULATORY PROTEIN).
GN   CDC24 OR CLS4 OR YAL041W.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87277425.
RA   MIYAMOTO S., OHYA Y., OHSUMI Y., ANRAKU Y.;
RL   GENE 54:125-132(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RX   MEDLINE; 95249563.
RA   BUSSEY H., KABACK D.B., ZHONG W., VO D.T., CLARK M.W., FORTIN N.,
RA   HALL J., OUELLETTE B.F.F., KENG T., BARTON A.B., SU Y., DAVIES C.K.,
```

```
RA   STORMS R.K.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 92:3809-3813(1995).
RN   [3]
RP   SIMILARITY TO CDC24 FAMILY.
RX   MEDLINE; 92095962.
RA   MIYAMOTO S., OHYA Y., SANO Y., SAKAGUCHI S., IIDA H., ANRAKU Y.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 181:604-610(1991).
CC   -!- FUNCTION: PROMOTES THE EXCHANGE OF CDC42-BOUND GDP BY GTP.
CC       CONTROL THE CALCIUM REGULATORY PROCESS OF BUD EMERGENCE. CDC24 MAY
CC       BE INVOLVED IN THE INITIAL SELECTION AND ORGANIZATION OF THE
CC       BUDDING SITE.
CC   -!- SIMILARITY: TO OTHER GUANINE-NUCLEOTIDE RELEASING FACTORS OF THE
CC       CDC24 FAMILY.
CC   -!- SIMILARITY: CONTAINS A PH DOMAIN.
DR   EMBL; M16809; G1100997; -.
DR   EMBL; U12980; G1101003; -.
DR   PIR; A27477; A27477.
DR   SGD; L0000262; CDC24.
DR   PROSITE; PS00741; GDS_CDC24; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   GUANINE-NUCLEOTIDE RELEASING FACTOR.
FT   DOMAIN      478    668       PH.
FT   DOMAIN      494    600       SER/THR-RICH.
FT   DOMAIN      681    778       SER/THR-RICH.
SQ   SEQUENCE   854 AA;  96939 MW;  E74FC7DC CRC32;

  Query Match           62.2%;  Score 51;  DB 1;  Length 854;
  Best Local Similarity  63.6%;  Pred. No. 5.14e+00;
  Matches     7;  Conservative    1;  Mismatches   3;  Indels   0;  Gaps   0;

Db    517 LFSEVVTKKSA 527
          ||  |:||| |
Qy      1 LFRAVITKKVA 11


RESULT   11
ID   BOB1_YEAST     STANDARD;      PRT;   980 AA.
AC   P38041;
DT   01-OCT-1994 (REL. 30, CREATED)
DT   01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   BOB1 PROTEIN (BEM1-BINDING PROTEIN).
GN   BOB1 OR BOI1 OR YBL085W OR YBL0717.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   BENDER A., BENDER L., KOKOJAN V.;
RL   SUBMITTED (APR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE; 96076635.
RA   OBERMAIER B., GASSENHUBER J., PIRAVANDI E., DOMDEY H.;
RL   YEAST 11:1103-1112(1995).
CC   -!- FUNCTION: BINDS TO THE BEM1 PROTEIN.
CC   -!- SIMILARITY: CONTAINS A COPY OF THE SH3 DOMAIN.
CC   -!- SIMILARITY: CONTAINS A PH DOMAIN.
DR   EMBL; L31406; G466436; -.
DR   EMBL; X79489; G496694; -.
DR   EMBL; Z35846; G536138; -.
DR   PIR; S45444; S45444.
DR   SGD; L0000191; BOI1.
DR   PROSITE; PS50002; SH3; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   SH3 DOMAIN.
FT   DOMAIN       13     77       SH3.
FT   DOMAIN      776    895       PH.
SQ   SEQUENCE   980 AA;  109295 MW;  99B7E197 CRC32;

  Query Match           62.2%;  Score 51;  DB 1;  Length 980;
  Best Local Similarity  54.5%;  Pred. No. 5.14e+00;
  Matches     6;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db     66 LYPAVFTKRIA 76
          |: ||:||::|
Qy      1 LFRAVITKKVA 11


RESULT   12
ID   RRP5_YEAST     STANDARD;      PRT;  1729 AA.
AC   Q05022;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   RRNA BIOGENESIS PROTEIN RRP5.
GN   RRP5 OR FMI1 OR YMR229C OR YM9959.11C.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   SKELTON J., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL   SUBMITTED (JUN-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE; 97051828.
RA   VENEMA J., TOLLERVEY D.;
RL   EMBO J. 15:5701-5714(1996).
CC   -!- FUNCTION: INVOLVED IN THE BIOGENESIS OF RRNA. REQUIRED FOR THE
CC       FORMATION OF 18S AND 5.8S RRNA.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
DR   EMBL; Z49939; G887610; -.
DR   SGD; L0003232; FMI1.
KW   NUCLEAR PROTEIN; RRNA PROCESSING.
SQ   SEQUENCE   1729 AA;  193133 MW;  F8D0B338 CRC32;

  Query Match           62.2%;  Score 51;  DB 1;  Length 1729;
  Best Local Similarity  54.5%;  Pred. No. 5.14e+00;
  Matches     6;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db   1674 LFERIITKKIT 1684
          ||  :||||::
Qy      1 LFRAVITKKVA 11


RESULT   13
ID   MAM2_SCHPO     STANDARD;      PRT;   348 AA.
AC   Q00619;
DT   01-APR-1993 (REL. 25, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   PHEROMONE P-FACTOR RECEPTOR.
GN   MAM2.
OS   SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=975;
RX   MEDLINE; 92037537.
RA   KITAMURA K., SHIMODA C.;
RL   EMBO J. 10:3743-3751(1991).
CC   -!- FUNCTION: RECEPTOR FOR THE PEPTIDE PHEROMONE P-FACTOR, A MATING
CC       FACTOR OF S.POMBE. PHEROMONE SIGNALING IS ESSENTIAL FOR INITIATION
CC       OF MEIOSIS IN S.POMBE; P-FACTOR SIGNALING ALONE MAY BE SUFFICIENT.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 4 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; X61672; G4978; -.
DR   PIR; S18521; S18521.
DR   GCRDB; GCR_0252; -.
KW   TRANSMEMBRANE; G-PROTEIN COUPLED RECEPTOR; PHEROMONE RESPONSE.
FT   TRANSMEM     46     69       POTENTIAL.
FT   TRANSMEM     79    103       POTENTIAL.
FT   TRANSMEM    125    141       POTENTIAL.
FT   TRANSMEM    162    180       POTENTIAL.
FT   TRANSMEM    207    225       POTENTIAL.
```

```
FT   TRANSMEM    249    267       POTENTIAL.
FT   TRANSMEM    283    301       POTENTIAL.
SQ   SEQUENCE   348 AA;  39285 MW;  68E94E15 CRC32;

  Query Match           61.0%;  Score 50;  DB 1;  Length 348;
  Best Local Similarity  55.6%;  Pred. No. 8.37e+00;
  Matches    5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    229 LFRAILIRK 237
          |||||:: :|
Qy      1 LFRAVITKK 9


RESULT   14
ID   FMDH_BACNO     STANDARD;      PRT;   394 AA.
AC   P17421;
DT   01-AUG-1990 (REL. 15, CREATED)
DT   01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT   01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE   POSSIBLE FIMBRIAL ASSEMBLY PROTEIN FIMD (SEROGROUP H1).
GN   FIMD.
OS   BACTEROIDES NODOSUS (DICHELOBACTER NODOSUS).
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; ANAEROBIC RODS;
OC   BACTEROIDACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SEROGROUP H1 ISOLATE VCS1215;
RX   MEDLINE; 91260439.
RA   HOBBS M., DALRYMPLE B.P., COX P.T., LIVINGSTONE S.P., DELANEY S.F.,
RA   MATTICK J.S.;
RL   MOL. MICROBIOL. 5:543-560(1991).
DR   EMBL; X52390; G580812; -.
DR   PIR; S15255; YQBZDH.
DR   PROSITE; PS00146; BETA_LACTAMASE_A; UNKNOWN_1.
KW   FIMBRIA.
SQ   SEQUENCE   394 AA;  45105 MW;  1310A727 CRC32;

  Query Match           61.0%;  Score 50;  DB 1;  Length 394;
  Best Local Similarity  72.7%;  Pred. No. 8.37e+00;
  Matches    8;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db    365 FRAASTKKTAD 375
          |||  ||| ||
Qy      2 FRAVITKKVAD 12


RESULT   15
ID   BPL1_YEAST     STANDARD;      PRT;   690 AA.
AC   P48445;
DT   01-FEB-1996 (REL. 33, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   BIOTIN--PROTEIN LIGASE (EC 6.3.4.-) (BIOTIN APO-PROTEIN LIGASE)
DE   (BIOTIN--[METHYLMALONYL-COA-CARBOXYLTRANSFERASE] LIGASE (EC 6.3.4.9) /
DE   BIOTIN--[PROPIONYL-COA-CARBOXYLASE (ATP-HYDROLYSING)] LIGASE
DE   (EC 6.3.4.10) (HOLOCARBOXYLASE SYNTHETASE) (HCS) / BIOTIN--
DE   [METHYLCROTONOYL-COA-CARBOXYLASE] LIGASE (EC 6.3.4.11) / BIOTIN--
DE   [ACETYL-COA-CARBOXYLASE] LIGASE (EC 6.3.4.15)).
GN   BPL1 OR ACC2 OR YDL141W OR D2140.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE; 95377607.
RA   CRONAN J.E. JR., WALLACE J.C.;
RL   FEMS MICROBIOL. LETT. 130:221-230(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RX   MEDLINE; 97127826.
RA   WOELFL S., HANEMAN V., SALUZ H.P.;
RL   YEAST' 12:1549-1554(1996).
RN   [3]
RP   SEQUENCE OF 1-291 FROM N.A.
RA   BARON L., LEGROS Y., BITEAU N., MONNET A., GRANOTIER C.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: POSTTRANSLATIONAL MODIFICATION OF SPECIFIC PROTEIN BY
CC       ATTACHMENT OF BIOTIN. ACTS ON VARIOUS CARBOXYLASES SUCH AS ACETYL-
CC       COA-CARBOXYLASE, PYRUVATE CARBOXYLASE, PROPIONYL COA CARBOXYLASE,
CC       AND 3-METHYLCROTONYL COA CARBOXYLASE.
CC   -!- CATALYTIC ACTIVITY: ATP + BIOTIN + APO-[METHYLMALONYL-COA:PYRUVATE
CC       CARBOXYLTRANSFERASE] = AMP + DIPHOSPHATE + [METHYLMALONYL-
CC       COA:PYRUVATE CARBOXYLTRANSFERASE].
CC   -!- CATALYTIC ACTIVITY: ATP + BIOTIN + APO-[PROPANOYL-COA:CARBON-
CC       DIOXIDE LIGASE (ADP-FORMING)] = AMP + DIPHOSPHATE + [PROPANOYL-
CC       COA:CARBON-DIOXIDE LIGASE (ADP-FORMING)].
CC   -!- CATALYTIC ACTIVITY: ATP + BIOTIN + APO-[3-METHYLCROTONOYL-
CC       COA:CARBON-DIOXIDE LIGASE (ADP-FORMING)] = AMP + DIPHOSPHATE + [3-
CC       METHYLCROTONOYL-COA:CARBON- DIOXIDE LIGASE (ADP-FORMING)].
CC   -!- CATALYTIC ACTIVITY: ATP + BIOTIN + APO-[ACETYL-COA:CARBON-DIOXIDE
CC       LIGASE (ADP FORMING)] = AMP + DIPHOSPHATE + [ACETYL-COA:CARBON-
CC       DIOXIDE LIGASE (ADP FORMING)].
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: WITH E.COLI BIRA AND OTHER EUKARYOTIC BIOTIN--PROTEIN
CC       LIGASES.
DR   EMBL; U27182; G886081; -.
DR   EMBL; X96876; E239063; -.
DR   EMBL; Z74189; E253238; -.
DR   SGD; L0002771; BPL1.
KW   LIGASE.
SQ   SEQUENCE   690 AA;  76363 MW;  AA6F29D3 CRC32;

  Query Match           61.0%;  Score 50;  DB 1;  Length 690;
  Best Local Similarity  50.0%;  Pred. No. 8.37e+00;
  Matches    5;  Conservative    5;  Mismatches   0;  Indels   0;  Gaps   0;

Db    679 IFKSLIAKKV 688
          :|:::|:|||
Qy      1 LFRAVITKKV 10


RESULT   16
ID   YOJN_ECOLI     STANDARD;      PRT;   890 AA.
AC   P39838; P47725; P47726; P76456;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   PROBABLE SENSOR PROTEIN YOJN (EC 2.7.3.-).
GN   YOJN.
OS   ESCHERICHIA COLI.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RA   BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL   SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE; 97251358.
RA   ITOH T., AIBA H., BABA T., FUJITA K., HAYASHI K., INADA T.,
RA   ISONO K., KASAI H., KIMURA S., KITAKAWA M., KITAGAWA M.,
RA   MAKINO K., MIKI T., MIZOBUCHI K., MORI H., MORI T., MOTOMURA K.,
RA   NAKADE S., NAKAMURA Y., NASHIMOTO H., NISHIO Y., OSHIMA T.,
RA   SAITO N., SAMPEI G., SEKI Y., SIVASUNDARAM S., TAGAMI H.,
RA   TAKEDA J., TAKEMOTO K., WADA C., YAMAMOTO Y., HORIUCHI T.;
RL   DNA RES. 3:379-392(1996).
RN   [3]
RP   SEQUENCE OF 1-596 FROM N.A.
RC   STRAIN=K12 / EMG2;
RA   ROBISON K., ESTEP P.E., O'KEEFFE T., CHURCH G.M.;
RL   SUBMITTED (OCT-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [4]
RP   SEQUENCE OF 593-890 FROM N.A.
```

```
RC   STRAIN=K12;
RX   MEDLINE; 90130299.
RA   STOUT V., GOTTESMAN S.;
RL   J. BACTERIOL. 172:659-669(1990).
RN   [5]
RP   IDENTIFICATION.
RX   MEDLINE; 95075659.
RA   BORODOVSKY M., RUDD K.E., KOONIN E.V.;
RL   NUCLEIC ACIDS RES. 22:4756-4767(1994).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC       (PROBABLE).
CC   -!- PTM: ACTIVATION PROBABLY REQUIRES A TRANSFER OF A PHOSPHATE GROUP
CC       BETWEEN A HIS IN THE TRANSMITTER DOMAIN AND A ASP OF THE RECEIVER
CC       DOMAIN (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC       KINASES.
CC   -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
CC       IN POSITIONS 145 AND 488.
CC   -!- CAUTION: REF.4 SEQUEMCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
CC       IN POSITIONS 597 AND 702.
DR   EMBL; AE000310; G1788545; -.
DR   EMBL; D90850; G1736857; -.
DR   EMBL; U38659; G1054951; ALT_FRAME.
DR   EMBL; U38659; G1054952; ALT_FRAME.
DR   EMBL; U38659; G1054953; ALT_FRAME.
DR   EMBL; M28242; -; NOT_ANNOTATED_CDS.
DR   ECOGENE; EG12385; YOJN.
KW   HYPOTHETICAL PROTEIN; SENSORY TRANSDUCTION; TRANSFERASE; KINASE;
KW   PHOSPHORYLATION; TRANSMEMBRANE; INNER MEMBRANE.
FT   DOMAIN        1     21       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     22     42       POTENTIAL.
FT   DOMAIN       43    308       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    309    329       POTENTIAL.
FT   DOMAIN      330    890       CYTOPLASMIC (POTENTIAL).
FT   CONFLICT     61     61       Q -> L (IN REF. 3).
FT   CONFLICT    164    164       K -> E (IN REF. 3).
FT   CONFLICT    174    174       T -> A (IN REF. 3).
FT   CONFLICT    176    176       S -> N (IN REF. 3).
FT   CONFLICT    179    179       V -> G (IN REF. 3).
FT   CONFLICT    184    184       A -> T (IN REF. 3).
FT   CONFLICT    267    267       T -> A (IN REF. 3).
FT   CONFLICT    598    599       ES -> DA (IN REF. 4).
SQ   SEQUENCE   890 AA;  100372 MW;  6D5BF6A8 CRC32;

  Query Match           61.0%;  Score 50;  DB 1;  Length 890;
  Best Local Similarity  60.0%;  Pred. No. 8.37e+00;
  Matches     6;  Conservative     4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    377 RTVISNKIAD 386
          |:||::|:||
Qy      3 RAVITKKVAD 12


RESULT   17
ID   PPS2_BACSU     STANDARD;      PRT;  2560 AA.
AC   P39846;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   PEPTIDE SYNTHETASE 2.
GN   PPSB OR PPS2.
OS   BACILLUS SUBTILIS.
OC   PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE; 95227362.
RA   TOGNONI A., FRANCHI E., MAGISTRELLI C., COLOMBO E., COSMINA P.,
RA   GRANDI G.;
RL   MICROBIOLOGY 141:645-648(1995).
CC   -!- COFACTOR: CONTAINS A COVALENTLY BOUND PHOSPHOPANTETHEINE
CC       (POTENTIAL).
CC   -!- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
CC       COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
DR   EMBL; Z34883; G509469; -.
DR   SUBTILIST; BG10971; PPSB.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR   PROSITE; PS00455; AMP_BINDING; 2.
DR   PROSITE; PS50075; ACP_DOMAIN; 2.
KW   MULTIFUNCTIONAL ENZYME; LIGASE; REPEAT; PHOSPHOPANTETHEINE.
FT   BINDING    2041   2041       PHOSPHOPANTETHEINE (POTENTIAL).
SQ   SEQUENCE   2560 AA;  290161 MW;  9BAA32F6 CRC32;

  Query Match           61.0%;  Score 50;  DB 1;  Length 2560;
  Best Local Similarity  41.7%;  Pred. No. 8.37e+00;
  Matches     5;  Conservative     5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     63 IFRTIFIKEVPD 74
          :||::: | |:|
Qy      1 LFRAVITKKVAD 12


RESULT   18
ID   VGLL_HSV2H     STANDARD;      PRT;   224 AA.
AC   P28278;
DT   01-DEC-1992 (REL. 24, CREATED)
DT   01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   GLYCOPROTEIN L PRECURSOR.
GN   GL OR UL1.
OS   HERPES SIMPLEX VIRUS (TYPE 2 / STRAIN HG52).
OC   VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92113549.
RA   MCGEOCH D.J., CUNNINGHAM C., MCINTYRE G., DOLAN A.;
RL   J. GEN. VIROL. 72:3057-3075(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   DOLAN A.;
RL   SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: ASSOCIATED WITH GLYCOPROTEIN H (GH) TO FORM A COMPLEX
CC       IMPORTANT FOR INFECTION AND CELL FUSION (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN L.
DR   EMBL; D10470; G221792; -.
DR   EMBL; Z86099; E304262; -.
DR   PIR; JQ1494; WMBEHG.
KW   GLYCOPROTEIN; SIGNAL.
FT   SIGNAL        1      ?       POTENTIAL.
FT   CHAIN         ?    224       GLYCOPROTEIN L.
FT   CARBOHYD    170    170       POTENTIAL.
SQ   SEQUENCE   224 AA;  25192 MW;  943BCE65 CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 224;
  Best Local Similarity  50.0%;  Pred. No. 1.35e+01;
  Matches     6;  Conservative     5;  Mismatches   1;  Indels   0;  Gaps   0;

Db     27 VLRSVIAKEVGD 38
          ::|:||:| |:|
Qy      1 LFRAVITKKVAD 12


RESULT   19
ID   DAPD_ACTPL     STANDARD;      PRT;   274 AA.
AC   P41396;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   2,3,4,5-TETRAHYDROPYRIDINE-2-CARBOXYLATE N-SUCCINYLTRANSFERASE
DE   (EC 2.3.1.117) (TETRAHYDRODIPICOLINATE N-SUCCINYLTRANSFERASE)
DE   (THP SUCCINYLTRANSFERASE) (TETRAHYDROPICOLINATE SUCCINYLASE).
GN   DAPD.
OS   ACTINOBACILLUS PLEUROPNEUMONIAE (HAEMOPHILUS PLEUROPNEUMONIAE).
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   PASTEURELLACEAE.
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=4074;
RX   MEDLINE; 94224145.
RA   LALONDE G., O'HANLEY P.D., STOCKER B.A., DENICH K.;
RL   MOL. MICROBIOL. 11:273-280(1994).
CC   -!- CATALYTIC ACTIVITY: SUCCINYL-COA + 2,3,4,5-TETRAHYDROPYRIDINE-
CC       2-CARBOXYLATE = COA + N-SUCCINYL-L-2-AMINO-6-OXOHEPTANEDIOATE.
CC   -!- PATHWAY: FOURTH STEP IN THE BIOSYNTHESIS OF DIAMINOPIMELATE AND
CC       LYSINE FROM ASPARTATE SEMIALDEHYDE.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC       ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
DR   EMBL; X63201; G38947; -.
DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW   TRANSFERASE; ACYLTRANSFERASE; REPEAT; LYSINE BIOSYNTHESIS;
KW   DIAMINOPIMELATE BIOSYNTHESIS.
SQ   SEQUENCE   274 AA;  29761 MW;  B0E49D1C CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 274;
  Best Local Similarity  70.0%;  Pred. No. 1.35e+01;
  Matches    7;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    246 LYCAVIVKKV 255
          |: ||| |||
Qy      1 LFRAVITKKV 10


RESULT   20
ID   DAPD_ECOLI     STANDARD;      PRT;   274 AA..
AC   P03948;
DT   23-OCT-1986 (REL. 02, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   2,3,4,5-TETRAHYDROPYRIDINE-2-CARBOXYLATE N-SUCCINYLTRANSFERASE
DE   (EC 2.3.1.117) (TETRAHYDRODIPICOLINATE N-SUCCINYLTRANSFERASE)
DE   (THP SUCCINYLTRANSFERASE) (TETRAHYDROPICOLINATE SUCCINYLASE).
GN   DAPD.
OS   ESCHERICHIA COLI.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 85054973.
RA   RICHAUD C., RICHAUD F., MARTIN C., HAZIZA C., PATTE J.-C.;
RL   J. BIOL. CHEM. 259:14824-14828(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / W3110;
RX   MEDLINE; 94261430.
RA   FUJITA N., MORI H., YURA T., ISHIHAMA A.;
RL   NUCLEIC ACIDS RES. 22:1637-1639(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RA   BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL   SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   SCHRAMM S., DUNCAN M., ALLEN E., ARAUJO R., APARICIO A., CHUNG E.,
RA   DAVIS K., FEDERSPIEL N., HYMAN R., KALMAN S., KOMP C., KURDI O.,
RA   LASHKARI D., LEW H., LIN D., NAMATH A., OEFNER P., ROBERTS D.,
RA   DAVIS R.W.;
RL   SUBMITTED (SEP-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [5]
RP   SEQUENCE OF 1-15 FROM N.A.
RC   STRAIN=K12 / W3110;
RX   MEDLINE; 94018640.
RA   VAN HEESWIJK W.C., RABENBERG M., WESTERHOFF H.V., KAHN D.D.;
RL   MOL. MICROBIOL. 9:443-458(1993).
RN   [6]
RP   SEQUENCE OF 1-11.
RC   STRAIN=K12 / W3110;
RA   PASQUALI C., SANCHEZ J.-C., RAVIER F., GOLAZ O., HUGHES G.J.,
RA   FRUTIGER S., PAQUET N., WILKINS M., APPEL R.D., BAIROCH A.,
RA   HOCHSTRASSER D.F.;
RL   SUBMITTED (SEP-1994) TO THE SWISS-PROT DATA BANK.
RN   [7]
RP   SEQUENCE OF 1-12.
RC   STRAIN=K12 / EMG2;
RA   LINK A.J.;
RL   SUBMITTED (OCT-1994) TO THE SWISS-PROT DATA BANK.
CC   -!- CATALYTIC ACTIVITY: SUCCINYL-COA + 2,3,4,5-TETRAHYDROPYRIDINE-
CC       2-CARBOXYLATE = COA + N-SUCCINYL-L-2-AMINO-6-OXOHEPTANEDIOATE.
CC   -!- PATHWAY: FOURTH STEP IN THE BIOSYNTHESIS OF DIAMINOPIMELATE AND
CC       LYSINE FROM ASPARTATE SEMIALDEHYDE.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC       ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
DR   EMBL; K02970; G145712; -.
DR   EMBL; D26562; G473821; -.
DR   EMBL; AE000126; G1786362; -.
DR   EMBL; U70214; G1552743; -.
DR   EMBL; Z21842; G49394; -.
DR   PIR; A00601; XNECSD.
DR   SWISS-2DPAGE; P03948; COLI.
DR   ECOGENE; EG10207; DAPD.
DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW   TRANSFERASE; ACYLTRANSFERASE; REPEAT; LYSINE BIOSYNTHESIS;
KW   DIAMINOPIMELATE BIOSYNTHESIS.
FT   CONFLICT     31     31       V -> D (IN REF. 1).
FT   CONFLICT    163    163       G -> R (IN REF. 1).
FT   CONFLICT    177    177       I -> M (IN REF. 1).
FT   CONFLICT    190    190       V -> L (IN REF. 1).
SQ   SEQUENCE   274 AA;  29892 MW;  4C9DF5CE CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 274;
  Best Local Similarity  70.0%;  Pred. No. 1.35e+01;
  Matches    7;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    247 LYCAVIVKKV 256
          |: ||| |||
Qy      1 LFRAVITKKV 10


RESULT   21
ID   DAPD_HAEIN     STANDARD;      PRT;   275 AA.
AC   P45284;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   2,3,4,5-TETRAHYDROPYRIDINE-2-CARBOXYLATE N-SUCCINYLTRANSFERASE
DE   (EC 2.3.1.117) (TETRAHYDRODIPICOLINATE N-SUCCINYLTRANSFERASE)
DE   (THP SUCCINYLTRANSFERASE) (TETRAHYDROPICOLINATE SUCCINYLASE).
GN   DAPD OR HI1634.
OS   HAEMOPHILUS INFLUENZAE.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   PASTEURELLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20;
RX   MEDLINE; 95350630.
RA   FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA   KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA   MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA   SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
RA   WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA   UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA   FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA   GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA   VENTER J.C.;
RL   SCIENCE 269:496-512(1995).
CC   -!- CATALYTIC ACTIVITY: SUCCINYL-COA + 2,3,4,5-TETRAHYDROPYRIDINE-
CC       2-CARBOXYLATE = COA + N-SUCCINYL-L-2-AMINO-6-OXOHEPTANEDIOATE.
CC   -!- PATHWAY: FOURTH STEP IN THE BIOSYNTHESIS OF DIAMINOPIMELATE AND
CC       LYSINE FROM ASPARTATE SEMIALDEHYDE.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
```

```
CC   -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC       ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
DR   EMBL; U32836; G1574480; ALT_INIT.
DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
DR   TIGR; HI1634; -.
KW   TRANSFERASE; ACYLTRANSFERASE; REPEAT; LYSINE BIOSYNTHESIS;
KW   DIAMINOPIMELATE BIOSYNTHESIS.
SQ   SEQUENCE   275 AA;  29723 MW;  E5F6024D CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 275;
  Best Local Similarity  70.0%;  Pred. No. 1.35e+01;
  Matches     7;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    247 LYCAVIVKKV 256
          |: ||| |||
Qy      1 LFRAVITKKV 10


RESULT   22
ID   GBA1_CRYNE     STANDARD;      PRT;   432 AA.
AC   P54853;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GUANINE NUCLEOTIDE-BINDING PROTEIN ALPHA SUBUNIT.
GN   GPA1.
OS   CRYPTOCOCCUS NEOFORMANS (FILOBASIDIELLA NEOFORMANS).
OC   EUKARYOTA; FUNGI; BASIDIOMYCOTINA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 42163;
RX   MEDLINE; 94274301.
RA   TOLKACHEVA T., MCNAMARA P., PIEKARZ E., COURCHESNE W.;
RL   INFECT. IMMUN. 62:2849-2856(1994).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS. INVOLVED IN THE MATING PATHWAY.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY.
DR   EMBL; U09372; G533774; -.
KW   GTP-BINDING; TRANSDUCER; MYRISTYLATION.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   LIPID         1      1       MYRISTATE (BY SIMILARITY).
FT   NP_BIND     118    125       GTP (BY SIMILARITY).
FT   NP_BIND     279    283       GTP (BY SIMILARITY).
FT   NP_BIND     348    351       GTP (BY SIMILARITY).
SQ   SEQUENCE   432 AA;  47848 MW;  7650F9E0 CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 432;
  Best Local Similarity  63.6%;  Pred. No. 1.35e+01;
  Matches     7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    147 FRGVIYKNVLD 157
          ||:|| |:| |
Qy      2 FRAVITKKVAD 12


RESULT   23
ID   ECHB_RAT       STANDARD;      PRT;   475 AA.
AC   Q60587;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MITOCHONDRIAL TRIFUNCTONAL ENZYME BETA SUBUNIT PRECURSOR (CONTAINS:
DE   3-KETOACYL-COA THIOLASE (EC 2.3.1.16) (ACETYL-COA ACYLTRANSFERASE)
DE   (BETA-KETOTHIOLASE)) (TP-BETA).
GN   HADHB.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR;
RX   MEDLINE; 94075334.
RA   KAMIJO T., AOYAMA T., MIYAZAKI J., HASHIMOTO T.;
RL   J. BIOL. CHEM. 268:26452-26460(1993).
CC   -!- CATALYTIC ACTIVITY: ACYL-COA + ACETYL-COA = COA + 3-OXOACYL-COA.
CC   -!- PATHWAY: THIRD STEP OF THE FATTY ACID BETA-OXIDATION CYCLE.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC   -!- SIMILARITY: BELONGS TO THE THIOLASE FAMILY.
DR   EMBL; D16479; G510110; -.
DR   PROSITE; PS00098; THIOLASE_1; 1.
DR   PROSITE; PS00099; THIOLASE_3; 1.
DR   PROSITE; PS00737; THIOLASE_2; 1.
KW   FATTY ACID METABOLISM; TRANSFERASE; ACYLTRANSFERASE; MITOCHONDRION;
KW   TRANSIT PEPTIDE.
FT   TRANSIT       1     34       MITOCHONDRION (POTENTIAL).
FT   CHAIN        35    475       3-KETOACYL-COA THIOLASE BETA-SUBUNIT OF
FT                                 TRIFUNCTIONAL PROTEIN.
FT   ACT_SITE    139    139       SUBSTRATE BINDING (BY SIMILARITY).
FT   ACT_SITE    459    459       BASE (BY SIMILARITY).
SQ   SEQUENCE   475 AA;  51414 MW;  BEBF9109 CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 475;
  Best Local Similarity  50.0%;  Pred. No. 1.35e+01;
  Matches     6;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db     89 LYRTNIPKDVVD 100
          |:|: | | | |
Qy      1 LFRAVITKKVAD 12


RESULT   24
ID   CTK1_YEAST     STANDARD;      PRT;   528 AA.
AC   Q03957;
DT   01-OCT-1993 (REL. 27, CREATED)
DT   01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   CTD KINASE ALPHA SUBUNIT (EC 2.7.1.-) (CTD KINASE 58 KD SUBUNIT)
DE   (CTDK-I ALPHA SUBUNIT).
GN   CTK1 OR YKL139W.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92314702.
RA   LEE J.M., GREENLEAF A.L.;
RL   GENE EXPR. 1:149-167(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   RAD M.R., XU G., KIRCHRATH L., FRITZ C., KEUCHEL H., HOLLENBERG C.P.;
RL   SUBMITTED (MAR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: CTDK-I HYPERPHOSPHORYLATES THE CARBOXYL-TERMINAL REPEAT
CC       DOMAIN (CTD) OF RNA POLYMERASE II LARGEST SUBUNIT. THIS PROTEIN IS
CC       THE CATALYTIC SUBUNIT.
CC   -!- SUBUNIT: CONSISTS OF THREE SUBUNITS (ALPHA, BETA, GAMMA) OF 58,
CC       38, AND 32 KD, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: BELONGS TO THE CDC2/CDC28 SUBFAMILY OF SER/THR
CC       PROTEIN KINASES.
DR   EMBL; M69024; G171328; -.
DR   EMBL; Z28139; G486235; -.
DR   PIR; S32593; S32593.
DR   SGD; L0000432; CTK1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW   NUCLEAR PROTEIN; PHOSPHORYLATION.
FT   DOMAIN       37     44       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      183    469       PROTEIN KINASE.
FT   NP_BIND     189    197       ATP (BY SIMILARITY).
FT   BINDING     212    212       ATP (BY SIMILARITY).
FT   ACT_SITE    306    306       BY SIMILARITY.
```

```
FT   DOMAIN      506    528       ASN/ASP-RICH.
FT   MOD_RES      14     14       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ   SEQUENCE   528 AA;  60501 MW;  4B629C3C CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 528;
  Best Local Similarity  66.7%;  Pred. No. 1.35e+01;
  Matches    6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db    162 RSVIAKKVP 170
          |:||:|||:
Qy      3 RAVITKKVA 11


RESULT   25
ID   KML2_CHICK     STANDARD;      PRT;  1906 AA.
AC   P19038;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   MYOSIN LIGHT CHAIN KINASE, NON-MUSCLE (EC 2.7.1.117).
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96033976.
RA   WATTERSON D.M., COLLINGE M., LUKAS T.J., VAN ELDIK L.J.,
RA   BIRUKOV K.G., STEPANOVA O.V., SHIRINSKY V.P.;
RL   FEBS LETT. 373:217-220(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (MLCK-108).
RX   MEDLINE; 90192792.
RA   OLSON N.J., PEARSON R.B., NEEDLEMAN D.S., HURWITZ M.J., KEMP B.E.,
RA   MEANS A.R.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 87:2284-2288(1990).
RN   [3]
RP   SEQUENCE OF 649-1906 FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=FIBROBLAST;
RX   MEDLINE; 90361738.
RA   SHOEMAKER M.O., LAU W., SHATTUCK R.L., KWIATKOWSKI A.P.,
RA   MATRISIAN P.E., GUERRA-SANTOS L., WILSON E., LUKAS T.J.,
RA   VAN ELDIK L.J., WATTERSON D.M.;
RL   J. CELL BIOL. 111:1107-1125(1990).
CC   -!- FUNCTION: MLCK PHOSPHORYLATES A SPECIFIC SERINE IN THE N-TERMINUS
CC       OF A MYOSIN LIGHT CHAIN, WHICH LEADS TO THE FORMATION CALMODULIN/
CC       MLCK SIGNAL TRANSDUCTION COMPLEXES WHICH ALLOW SELECTIVE
CC       TRANSDUCTION OF CALCIUM SIGNALS.
CC   -!- CATALYTIC ACTIVITY: ATP + [MYOSIN LIGHT-CHAIN] = ADP + [MYOSIN
CC       LIGHT-CHAIN] PHOSPHATE.
CC   -!- ALTERNATIVE PRODUCTS: A SINGLE NUCLEAR GENE PRODUCES BOTH FORMS
CC       BY USE OF ALTERNATIVE INITIATION CODONS IN THE SAME READING
CC       FRAME.
DR   EMBL; X52876; G992993; -.
DR   EMBL; X52876; G992994; -.
DR   EMBL; M31048; G212661; -.
DR   PIR; A37099; A37099.
DR   PIR; S11652; S11652.
DR   HSSP; P02593; 1CDL.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; CALMODULIN-BINDING;
KW   ATP-BINDING; PHOSPHORYLATION; REPEAT; ALTERNATIVE INITIATION.
FT   CHAIN         1   1906       MLCK-210.
FT   CHAIN       935   1906       MLCK-108.
FT   DOMAIN     1453   1708       PROTEIN KINASE.
FT   NP_BIND    1459   1467       ATP (BY SIMILARITY).
FT   BINDING    1482   1482       ATP (BY SIMILARITY).
FT   ACT_SITE   1574   1574       BY SIMILARITY.
FT   DOMAIN     1716   1728       CALMODULIN AUTOINHIBITION (AM13) REGION
FT                                (POTENTIAL).
FT   DOMAIN     1730   1749       CALMODULIN RECOGNITION (RS20) REGION
FT                                (POTENTIAL).
FT   DOMAIN     1317   1364       MOTIF IA.
FT   DOMAIN     1385   1402       MOTIF IB.
FT   DOMAIN      660   1833       4 X REPEATS, MOTIF IIA.
FT   REPEAT      660    676       IIA-1.
FT   REPEAT      758    774       IIA-2.
FT   REPEAT     1107   1123       IIA-3.
FT   REPEAT     1817   1833       IIA-4.
FT   DOMAIN      693   1866       5 X REPEATS, MOTIF IIB.
FT   REPEAT      693    708       IIB-1.
FT   REPEAT      791    807       IIB-2.
FT   REPEAT     1140   1156       IIB-3.
FT   REPEAT     1281   1297       IIB-4.
FT   REPEAT     1851   1866       IIB-5.
FT   DOMAIN      970   1226       4 X REPEATS, MOTIF III.
FT   REPEAT      970    987       III-1.
FT   REPEAT      999   1016       III-2.
FT   REPEAT     1061   1078       III-3.
FT   REPEAT     1209   1226       III-4.
FT   MOD_RES    1748   1748       PHOSPHORYLATION.
FT   MOD_RES    1762   1762       PHOSPHORYLATION.
SQ   SEQUENCE   1906 AA;  210445 MW;  0D2515AE CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 1906;
  Best Local Similarity  50.0%;  Pred. No. 1.35e+01;
  Matches    5;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    909 FRDILGKKVS 918
          || :: |||:
Qy      2 FRAVITKKVA 11


RESULT   26
ID   Y226_METJA     STANDARD;      PRT;   193 AA.
AC   Q57679;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL PROTEIN MJ0226.
GN   MJ0226.
OS   METHANOCOCCUS JANNASCHII.
OC   ARCHAEBACTERIA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96337999.
RA   BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA   SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA   KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA   OVERBEEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA   SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA   UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA   COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA   KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RL   SCIENCE 273:1058-1073(1996).
CC   -!- SIMILARITY: TO YEAST HAM1.
DR   EMBL; U67478; G1590963; -.
DR   TIGR; MJ0226; -.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   193 AA;  22202 MW;  DF81189B CRC32;

  Query Match           58.5%;  Score 48;  DB 1;  Length 193;
  Best Local Similarity  25.0%;  Pred. No. 2.16e+01;
  Matches    3;  Conservative    7;  Mismatches   2;  Indels   0;  Gaps   0;

Db    129 LFKGIVKGRVSE 140
          ||:::: :|::
Qy      1 LFRAVITKKVAD 12


RESULT   27
ID   FLIP_CAUCR     STANDARD;      PRT;   266 AA.
AC   Q45980;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
```

```
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   FLAGELLAR BIOSYNTHETIC PROTEIN FLIP.
GN   FLIP.
OS   CAULOBACTER CRESCENTUS.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; BUDDING AND/OR APPENDAGED.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CB15;
RX   MEDLINE; 95325304.
RA   GOBER J.W., BOYD C.H., JARVIS M., MANGAN E.K., RIZZO M.F.,
RA   WINGROVE J.A.;
RL   J. BACTERIOL. 177:3656-3667(1995).
CC   -!- FUNCTION: PLAYS A ROLE IN THE FLAGELLUM-SPECIFIC TRANSPORT SYSTEM
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE FLIP/MOPC/SPAP FAMILY.
DR   EMBL; U20387; G841253; -.
DR   PROSITE; PS01060; FLIP_1; 1.
DR   PROSITE; PS01061; FLIP_2; 1.
KW   FLAGELLA; TRANSMEMBRANE.
FT   TRANSMEM     20     40       POTENTIAL.
FT   TRANSMEM     58     78       POTENTIAL.
FT   TRANSMEM    102    122       POTENTIAL.
FT   TRANSMEM    202    222       POTENTIAL.
FT   TRANSMEM    226    246       POTENTIAL.
SQ   SEQUENCE   266 AA;  28527 MW;  C506FCD8 CRC32;

  Query Match           58.5%;  Score 48;  DB 1;  Length 266;
  Best Local Similarity  50.0%;  Pred. No. 2.16e+01;
  Matches     6;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db       5 LFKSVIGAKAED 16
           ||::||  |  |
Qy       1 LFRAVITKKVAD 12


RESULT   28
ID   THI4_FUSOX     STANDARD;      PRT;   320 AA.
AC   P23618;
DT   01-NOV-1991 (REL. 20, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   THIAZOLE BIOSYNTHETIC ENZYME (STRESS-INDUCIBLE PROTEIN STI35).
GN   STI35.
OS   FUSARIUM OXYSPORUM.
OC   EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SP. CUCUMERINUM OWEN, ISOLATE B1-GK;
RX   MEDLINE; 90330561.
RA   CHOI G.H., MAREK E.T., SCHARDL C.L., RICHEY M.G., CHANG S.,
RA   SMITH D.A.;
RL   J. BACTERIOL. 172:4522-4528(1990).
CC   -!- FUNCTION: INVOLVED IN BIOSYNTHESIS OF THE THIAMINE PRECURSOR
CC       THIAZOLE (BY SIMILARITY).
CC   -!- INDUCTION: BY ETHANOL, COPPER(II) CHLORIDE AND HEAT.
CC   -!- SIMILARITY: BELONGS TO THE THI4 FAMILY OF ENZYMES.
DR   EMBL; M33643; G168164; -.
DR   PIR; B37767; B37767.
KW   THIAMIN BIOSYNTHESIS; HEAT SHOCK.
SQ   SEQUENCE   320 AA;  34579 MW;  94180D07 CRC32;

  Query Match           58.5%;  Score 48;  DB 1;  Length 320;
  Best Local Similarity  58.3%;  Pred. No. 2.16e+01;
  Matches     7;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db     118 LFSAMIMRKPAD 129
           || |:| :| ||
Qy       1 LFRAVITKKVAD 12


RESULT   29
ID   B3AR_CAVPO     STANDARD;      PRT;   351 AA.
AC   Q60483;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   BETA-3 ADRENERGIC RECEPTOR (FRAGMENT).
GN   ADRB3.
OS   CAVIA PORCELLUS (GUINEA PIG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 97151378.
RA   ATGIE C., TAVERNIER G., D'ALLAIRE F., BENGTSSON T., MARTI L.,
RA   CARPENE C., LAFONTAN M., BUKOWIECKI L.J., LANGIN D.;
RL   AM. J. PHYSIOL. 271:R1729-R1738(1996).
CC   -!- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC       INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC       PROTEINS. BETA-3 IS INVOLVED IN THE REGULATION OF LIPOLYSIS AND
CC       THERMOGENESIS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: WHITE AND BROWN ADIPOSE TISSUES.
CC   -!- THE GUINEA PIG DIFFERS FROM OTHER RODENTS BY AN ABSENCE OF BETA-3
CC       ADRENERGIC EFFECTS AND BY LOW EXPRESSION IN BROWN AND WHITE
CC       ADIPOSE TISSUES. IT IS CLOSER TO HUMAN OR PRIMATE THAN RODENT
CC       BETA-3.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; U51098; G1256416; -.
DR   PROSITE; PS00237; G_PROTEIN_RECEPTOR; 1.
KW   G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW   MULTIGENE FAMILY.
FT   DOMAIN        1     36       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     37     60       1 (POTENTIAL).
FT   DOMAIN       61     69       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     70     88       2 (POTENTIAL).
FT   DOMAIN       89    108       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    109    130       3 (POTENTIAL).
FT   DOMAIN      131    152       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    153    175       4 (POTENTIAL).
FT   DOMAIN      176    200       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    201    222       5 (POTENTIAL).
FT   DOMAIN      223    290       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    291    312       6 (POTENTIAL).
FT   DOMAIN      313    324       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    325    345       7 (POTENTIAL).
FT   DOMAIN      346   >351       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    107    186       BY SIMILARITY.
FT   CARBOHYD      8      8       POTENTIAL.
FT   CARBOHYD     26     26       POTENTIAL.
FT   NON_TER     351    351
SQ   SEQUENCE   351 AA;  37364 MW;  72E17433 CRC32;

  Query Match           58.5%;  Score 48;  DB 1;  Length 351;
  Best Local Similarity  62.5%;  Pred. No. 2.16e+01;
  Matches     5;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db     142 YRAVVTKR 149
           :|||:||:
Qy       2 FRAVITKK 9


RESULT   30
ID   FADH_METMR     STANDARD;      PRT;   424 AA.
AC   P47734;
DT   01-FEB-1996 (REL. 33, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   GLUTATHIONE-DEPENDENT FORMALDEHYDE DEHYDROGENASE (EC 1.2.1.1) (FDH)
DE   (FALDH).
GN   FDH.
OS   METHYLOBACTER MARINUS.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   METHYLOCOCCACEAE.
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A45;
RX   MEDLINE; 95011501.
RA   SPEER B.S., CHISTOSERDOVA L., LIDSTROM M.E.;
RL   FEMS MICROBIOL. LETT. 121:349-355(1994).
CC   -!- CATALYTIC ACTIVITY: FORMALDEHYDE + GLUTATHIONE + NAD(+) =
CC       S-FORMYLGLUTATHIONE + NADH.
CC   -!- COFACTOR: REQUIRES ZINC FOR ITS ACTIVITY (POTENTIAL).
CC   -!- COFACTOR: EQUALLY ACTIVE ON NAD AND NADP.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC       FAMILY. BELONGS TO THE ADH CLASS-III SUBFAMILY.
DR   EMBL; L33464; G496118; -.
DR   PROSITE; PS00059; ADH_ZINC; 1.
KW   OXIDOREDUCTASE; ZINC; NAD; NADP.
FT   METAL        72     72       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL        94     94       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       124    124       ZINC (SECOND ATOM) (BY SIMILARITY).
FT   METAL       127    127       ZINC (SECOND ATOM) (BY SIMILARITY).
FT   METAL       130    130       ZINC (SECOND ATOM) (BY SIMILARITY).
FT   METAL       138    138       ZINC (SECOND ATOM) (BY SIMILARITY).
SQ   SEQUENCE   424 AA;  46033 MW;  5DAB3AD0 CRC32;

  Query Match           58.5%;  Score 48;  DB 1;  Length 424;
  Best Local Similarity  62.5%;  Pred. No. 2.16e+01;
  Matches     5;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db     29 LFRSVVMK 36
          |||:|: |
Qy      1 LFRAVITK 8


RESULT   31
ID   RSP3_CHLRE     STANDARD;      PRT;   516 AA.
AC   P12759;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE   RADIAL SPOKE PROTEIN 3.
GN   RSP3.
OS   CHLAMYDOMONAS REINHARDTII.
OC   EUKARYOTA; PLANTA; PHYCOPHYTA; CHLOROPHYTA (GREEN ALGAE);
OC   CHLOROPHYCEAE; VOLVOCALES; CHLAMYDOMONADACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=21GR;
RX   MEDLINE; 89308863.
RA   WILLIAMS B.D., VELLECA M.A., CURRY A.M., ROSENBAUM J.L.;
RL   J. CELL BIOL. 109:235-245(1989).
CC   -!- FUNCTION: PROTEIN 3 MAY ATTACH THE RADIAL SPOKE TO THE OUTER
CC       DOUBLET MICROTUBULE OR IS REQUIRED TO FORM A STABLE SPOKE
CC       STRUCTURE.
CC   -!- FUNCTION: FLAGELLAR RADIAL SPOKES CONTRIBUTE TO THE REGULATION
CC       OF DYNEIN ARM ACTIVITY AND THUS THE PATTERN OF FLAGELLAR BENDING.
CC       THEY CONSIST OF A THIN STALK, WHICH IS ATTACHED TO THE A SUBFIBER
CC       OF THE OUTER DOUBLET MICROTUBULE, AND A BULBOUS HEAD, WHICH IS
CC       ATTACHED TO THE STALK AND APPEARS TO INTERACT WITH THE
CC       PROJECTIONS FROM THE CENTRAL PAIR OF MICROTUBULES.
CC   -!- SUBCELLULAR LOCATION: RADIAL SPOKE.
CC   -!- PTM: PROTEIN 3 IS ONE OF THE 5 RADIAL SPOKE PROTEINS THAT ARE
CC       PHOSPHORYLATED.
CC   -!- PTM: THE ISOFORM PROTEIN 3A MIGHT BE SIMPLY THE UNPHOSPHORYLATED
CC       POLYPEPTIDE.
DR   EMBL; X14549; G18218; -.
DR   PIR; S05962; S05962.
DR   PIR; A31270; A31270.
KW   FLAGELLA; PHOSPHORYLATION.
SQ   SEQUENCE   516 AA;  56784 MW;  9768D176 CRC32;

  Query Match           58.5%;  Score 48;  DB 1;  Length 516;
  Best Local Similarity  41.7%;  Pred. No. 2.16e+01;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db    290 LARGVVARRVVD 301
          | |:|::::| |
Qy      1 LFRAVITKKVAD 12


RESULT   32
ID   AMYM_BACLI     STANDARD;      PRT;   578 AA.
AC   Q04977;
DT   01-OCT-1994 (REL. 30, CREATED)
DT   01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   MALTOGENIC ALPHA-AMYLASE PRECURSOR (EC 3.2.1.133) (GLUCAN 1,4-ALPHA-
DE   MALTOHYDROLASE).
GN   BLMA.
OS   BACILLUS LICHENIFORMIS.
OC   PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 27811;
RX   MEDLINE; 93054487.
RA   KIM I.C., CHA J.H., KIM J.R., JANG S.Y., SEO B.C., CHEONG T.K.,
RA   LEE D.S., CHOI Y.D., PARK K.H.;
RL   J. BIOL. CHEM. 267:22108-22114(1992).
CC   -!- FUNCTION: CONVERTS STARCH INTO MALTOSE. IN CONTRARY TO OTHER
CC       MALTOGENIC ALPHA-AMYLASES BLMA CANNOT HYDROLYSE 1,4-ALPHA-
CC       GLUCOSIDIC LINKAGE NEXT TO 1,6-ALPHA-GLUCOSIDIC LINKAGES.
CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-ALPHA-D-GLUCOSIDIC LINKAGES
CC       IN POLYSACCHARIDES SO AS TO REMOVE SUCCESSIVE ALPHA-MALTOSE UNITS
CC       FROM THE NON-REDUCING ENDS OF THE CHAINS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR   EMBL; X67133; G39564; -.
DR   PIR; A44326; A44326.
KW   HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL.
FT   SIGNAL        1      ?       POTENTIAL.
FT   CHAIN         ?    578       MALTOGENIC ALPHA-AMYLASE.
SQ   SEQUENCE   578 AA;  66924 MW;  E5888900 CRC32;

  Query Match           58.5%;  Score 48;  DB 1;  Length 578;
  Best Local Similarity  50.0%;  Pred. No. 2.16e+01;
  Matches     4;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    221 LFRTVVSR 228
          |||:|:::
Qy      1 LFRAVITK 8


RESULT   33
ID   YAJ3_SCHPO     STANDARD;      PRT;   754 AA.
AC   Q09903;
DT   01-FEB-1996 (REL. 33, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   PUTATIVE ATP-DEPENDENT RNA HELICASE C30D11.03.
GN   SPAC30D11.03.
OS   SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   PEARSON D., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL   SUBMITTED (NOV-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE.
CC   -!- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES.
DR   EMBL; Z67961; E209203; -.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   HYPOTHETICAL PROTEIN; ATP-BINDING; RNA-BINDING; HELICASE.
FT   NP_BIND     303    310       ATP (POTENTIAL).
FT   SITE        412    415       DEAD BOX.
SQ   SEQUENCE   754 AA;  85452 MW;  5D3AA652 CRC32;

  Query Match           58.5%;  Score 48;  DB 1;  Length 754;
```

```
  Best Local Similarity  58.3%;  Pred. No. 2.16e+01;
  Matches     7;  Conservative    1;  Mismatches   4;  Indels   0;  Gaps   0;

Db    441 LFSATMTDKVDD 452
          || | :| || |
Qy      1 LFRAVITKKVAD 12


RESULT   34
ID   RM39_YEAST     STANDARD;      PRT;    69 AA.
AC   P36533;
DT   01-JUN-1994 (REL. 29, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   MITOCHONDRIAL 60S RIBOSOMAL PROTEIN L39 (YML39).
GN   MRPL39 OR YML009C OR YM9571.09C.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   GENTLES S., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL   SUBMITTED (JUN-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE OF 1-49.
RX   MEDLINE; 91285106.
RA   GROHMANN L., GRAACK H.-R., KRUFT V., CHOLI T., GOLDSCHMIDT-REISIN S.,
RA   KITAKAWA M.;
RL   FEBS LETT. 284:51-56(1991).
DR   EMBL; Z49810; G854481; -.
DR   PIR; S17278; S17278.
DR   SGD; L0002695; MRPL39.
KW   RIBOSOMAL PROTEIN; MITOCHONDRION.
FT   INIT_MET      0      0
FT   CONFLICT     28     28       K -> R (IN REF. 2).
FT   CONFLICT     37     37       R -> K (IN REF. 2).
FT   CONFLICT     46     46       V -> R (IN REF. 2).
SQ   SEQUENCE   69 AA;  7841 MW;  3BAFE480 CRC32;

  Query Match           57.3%;  Score 47;  DB 1;  Length 69;
  Best Local Similarity  41.7%;  Pred. No. 3.43e+01;
  Matches     5;  Conservative    3;  Mismatches   4;  Indels   0;  Gaps   0;

Db     47 LFKEAKKRKVAE 58
          ||:    :|||:
Qy      1 LFRAVITKKVAD 12


RESULT   35
ID   RL9_HAEIN      STANDARD;      PRT;   149 AA.
AC   P44349;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   50S RIBOSOMAL PROTEIN L9.
GN   RPLI OR RPL9 OR HI0544.
OS   HAEMOPHILUS INFLUENZAE.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   PASTEURELLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20;
RX   MEDLINE; 95350630.
RA   FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA   KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA   MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA   SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
RA   WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA   UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA   FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA   GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA   VENTER J.C.;
RL   SCIENCE 269:496-512(1995).
CC   -!- FUNCTION: BINDS TO THE 23S RRNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L9P FAMILY OF RIBOSOMAL PROTEINS.
DR   EMBL; U32736; G1573529; -.
DR   PROSITE; PS00651; RIBOSOMAL_L9; 1.
DR   TIGR; HI0544; -.
KW   RIBOSOMAL PROTEIN; RRNA-BINDING.
SQ   SEQUENCE   149 AA;  15636 MW;  30C08516 CRC32;

  Query Match           57.3%;  Score 47;  DB 1;  Length 149;
  Best Local Similarity  50.0%;  Pred. No. 3.43e+01;
  Matches     6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db     90 LFGAITTRDVAE 101
          || |: |: ||:
Qy      1 LFRAVITKKVAD 12


Search completed: Tue Apr  7 08:39:41 1998
Job time : 9 secs.
```

****

MPSRCH (TM)

****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:39:59 1998; MasPar time 4.83 Seconds
104.542 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-3
Description: (1-12) from 5541104.pep
Perfect Score: 82
Sequence: 1 LFRAVITKKVAD 12

Scoring table: PAM 150
Gap 15

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: sptrembl5
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phage 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 25.051; Variance 28.357; scale 0.883

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	61	74.4	953	3	Q19753	F23B12.7.	4.55e-02
2	55	67.1	875	9	P95949	ATP-DEPENDENT HELICASE	1.09e+00
3	54	65.9	693	12	Q91889	UBIQUITIN-LIKE FUSION	1.82e+00
4	54	65.9	701	12	Q91890	UBIQUITIN-LIKE FUSION	1.82e+00
5	53	64.6	362	3	Q22897	COSMID C16D9.	3.00e+00
6	52	63.4	317	2	Q14798	MAGE-4 PROTEIN.	4.92e+00
7	52	63.4	663	3	O16057	LIM DOMAIN PROTEIN.	4.92e+00
8	51	62.2	380	3	Q20049	F35G12.2.	8.02e+00
9	49	59.8	201	8	O23469	SIMILARITY TO MYROSINA	2.08e+01
10	49	59.8	224	11	Q69112	ULLORF.	2.08e+01
11	49	59.8	236	9	Q58415	HYPOTHETICAL PROTEIN M	2.08e+01
12	49	59.8	287	9	O24963	HYPOTHETICAL 32.9 KD P	2.08e+01
13	49	59.8	446	3	P91320	SIMILAR TO ACETYLCHOLI	2.08e+01
14	49	59.8	447	3	Q23501	ZK455.8.	2.08e+01
15	49	59.8	481	3	Q19905	F29F11.1.	2.08e+01
16	49	59.8	522	2	Q92573	MYELOBLAST KIAA0241 (F	2.08e+01
17	49	59.8	660	9	Q45923	SIMILAR TO TNPB OF TRA	2.08e+01
18	49	59.8	813	10	Q63315	LONG TYPE PB-CADHERIN.	2.08e+01
19	49	59.8	851	8	O23524	HYPOTHETICAL 96.8 KD P	2.08e+01
20	49	59.8	1454	1	Q06164	CHROMOSOME XII COSMID	2.08e+01
21	49	59.8	1509	10	Q61194	P170 PHOSPHATIDYLINOSI	2.08e+01
22	49	59.8	1658	10	Q61182	PHOSPHOINOSITIDE 3-KIN	2.08e+01
23	49	59.8	1686	2	O00443	PHOSPHOINOSITIDE 3-KIN	2.08e+01
24	49	59.8	1807	1	O13661	HYPOTHETICAL 229.9KD P	2.08e+01
25	48	58.5	131	12	Q98914	CHICK OLFACTORY RECEPT	3.31e+01
26	48	58.5	133	9	Q53037	4.6KDA PROTEIN.	3.31e+01
27	48	58.5	167	11	P87594	P346, 54-KDA RNA DEPEN	3.31e+01
28	48	58.5	167	11	Q66247	P18, P13, P20, AND P23	3.31e+01
29	48	58.5	281	6	O21267	ATP SYNTHASE F1 SUBUNI	3.31e+01
30	48	58.5	319	9	O28219	ACTIVATOR 1, REPLICATI	3.31e+01
31	48	58.5	336	9	Q58293	HYPOTHETICAL PROTEIN M	3.31e+01
32	48	58.5	458	9	O26038	FERRODOXIN-LIKE PROTEI	3.31e+01
33	48	58.5	478	9	O32912	INOSINE-5'-MONOPHOSPHA	3.31e+01
34	48	58.5	624	8	O04130	PHOSPHOGLYCERATE DEHYD	3.31e+01
35	48	58.5	793	9	P73465	DNA HELICASE II.	3.31e+01
36	48	58.5	1253	3	Q19523	F17C8.1.	3.31e+01
37	48	58.5	2165	11	O09721	POLYMERASE (L).	3.31e+01
38	48	58.5	2165	11	Q82027	RNA POLYMERASE LARGE S	3.31e+01
39	47	57.3	87	9	O27942	CONSERVED HYPOTHETICAL	5.23e+01
40	47	57.3	94	10	Q63229	GUANINE NUCLEOTIDE REG	5.23e+01
41	47	57.3	266	3	Q18669	HYPOTHETICAL PROTEIN C	5.23e+01
42	47	57.3	268	9	Q58489	HYPOTHETICAL PROTEIN M	5.23e+01
43	47	57.3	309	3	Q27083	CLOTTING FACTOR G BETA	5.23e+01
44	47	57.3	353	3	O15975	GTP-BINDING PROTEIN Q	5.23e+01
45	47	57.3	355	3	Q17386	EGL-30.	5.23e+01
46	47	57.3	359	2	O15108	GTP-BINDING PROTEIN AL	5.23e+01
47	47	57.3	359	2	O15109	GUANINE NUCLEOTIDE BIN	5.23e+01
48	47	57.3	359	10	Q61939	GUANINE NUCLEOTIDE BIN	5.23e+01
49	47	57.3	367	10	Q60942	POTASSIUM VOLTAGE GATE	5.23e+01
50	47	57.3	399	9	O27322	ARGININOSUCCINATE SYNT	5.23e+01
51	47	57.3	437	2	Q99764	HYPOTHETICAL 48.5 KD P	5.23e+01
52	47	57.3	457	3	O01907	SIMILAR TO ALPHA TUBUL	5.23e+01
53	47	57.3	610	2	Q92732	RNA HELICASE.	5.23e+01
54	47	57.3	627	9	Q59455	PYRUVATE SYNTHASE (EC	5.23e+01
55	47	57.3	679	10	Q62430	MMMRE11B.	5.23e+01
56	47	57.3	706	10	Q61216	PUTATIVE ENDO/EXONUCLE	5.23e+01
57	47	57.3	722	10	O08679	SERINE/THREONINE KINAS	5.23e+01
58	47	57.3	745	2	Q15524	SERINE/THREONINE PROTE	5.23e+01
59	47	57.3	991	2	Q14844	MYOSIN LIGHT CHAIN KIN	5.23e+01
60	47	57.3	998	11	Q66929	PROTEIN A.	5.23e+01
61	47	57.3	1037	2	O15397	RANBP8.	5.23e+01
62	47	57.3	1147	4	Q28729	SMOOTH MUSCLE MYOSIN L	5.23e+01
63	47	57.3	1176	4	Q28824	155 KDA MYOSIN LIGHT C	5.23e+01
64	47	57.3	1687	10	O35651	DHM2 PROTEIN.	5.23e+01
65	47	57.3	1706	10	P97790	5'-3' EXONUCLEASE.	5.23e+01
66	47	57.3	1719	10	P97789	5'-3' EXONUCLEASE.	5.23e+01
67	47	57.3	1914	2	Q15746	MYOSIN LIGHT CHAIN KIN	5.23e+01
68	46	56.1	112	9	Q58983	HYPOTHETICAL 12.8 KD P	8.19e+01
69	46	56.1	123	1	P87267	YDR417CP.	8.19e+01
70	46	56.1	131	3	O18236	Y57G11C.12.	8.19e+01
71	46	56.1	175	9	O07562	HYPOTHETICAL 20.4 KD P	8.19e+01
72	46	56.1	215	9	Q59805	CHLORAMPHENICOL ACETYL	8.19e+01
73	46	56.1	215	9	Q59436	CHLORAMPHENICOL ACETYL	8.19e+01
74	46	56.1	237	3	O17116	F39E9.1 PROTEIN.	8.19e+01
75	46	56.1	238	3	Q18773	SIMILAR TO RAS-RELATED	8.19e+01
76	46	56.1	273	1	Q07505	CHROMOSOME IV READING	8.19e+01
77	46	56.1	279	8	O04361	PROHIBITIN.	8.19e+01
78	46	56.1	330	9	Q45263	NODULATION PROTEIN.	8.19e+01
79	46	56.1	331	9	Q44602	PHOSPHORIBOSYL ANTHRAN	8.19e+01
80	46	56.1	346	8	Q41641	FERREDOXIN-NADP+ REDUC	8.19e+01
81	46	56.1	350	1	Q03130	O3635P (FRAGMENT).	8.19e+01
82	46	56.1	357	9	O29920	CONSERVED HYPOTHETICAL	8.19e+01
83	46	56.1	412	8	Q42761	SQUALENE SYNTHASE (EC	8.19e+01
84	46	56.1	475	4	O18757	PEROXISOMAL CA-DEPENDE	8.19e+01
85	46	56.1	482	2	Q92879	CUG-BP/HNAB50.	8.19e+01
86	46	56.1	509	9	Q58485	HYPOTHETICAL PROTEIN M	8.19e+01
87	46	56.1	551	10	P70459	ETS-DOMAIN TRANSCRIPTI	8.19e+01
88	46	56.1	658	11	O41288	GAG PROTEIN.	8.19e+01
89	46	56.1	696	9	O07651	FERRIC RECEPTOR CFRA.	8.19e+01
90	46	56.1	728	9	Q55262	GLYCOSYLTREHALOSE-PROD	8.19e+01
91	46	56.1	732	9	P95869	ALPHA-AMYLASE.	8.19e+01
92	46	56.1	907	12	Q98850	NEURONAL MYOSIN LIGHT	8.19e+01
93	46	56.1	1203	3	O17720	C55A6.2.	8.19e+01

94	46	56.1	1242	3	Q18033	COSMID C16A3.	8.19e+01
95	46	56.1	1265	10	Q61992	RAN BINDING PROTEIN 2	8.19e+01
96	46	56.1	1283	3	Q94887	NEUREXIN IV PRECURSOR.	8.19e+01
97	46	56.1	1313	1	P87141	HYPOTHETICAL 148.5 KD	8.19e+01
98	46	56.1	1365	11	O37174	REPLICASE.	8.19e+01
99	46	56.1	1365	11	Q65005	ORF1=155K.	8.19e+01
100	46	56.1	4131	3	Q19542	F18C12.1.	8.19e+01

ALIGNMENTS

```
RESULT    1
ID   Q19753         PRELIMINARY;   PRT;   953 AA.
AC   Q19753;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   F23B12.7.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   WILD A.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
DR   EMBL; Z77659; E255897; -.
SQ   SEQUENCE   953 AA;  108423 MW;  080348D8 CRC32;

  Query Match           74.4%;  Score 61;  DB 3;  Length 953;
  Best Local Similarity  41.7%;  Pred. No. 4.55e-02;
  Matches    5;  Conservative    6;  Mismatches   1;  Indels   0;  Gaps   0;

Db    388 LFKTIVGKKITD 399
          ||:::: ||::|
Qy      1 LFRAVITKKVAD 12


RESULT    2
ID   P95949         PRELIMINARY;   PRT;   875 AA.
AC   P95949;
DT   01-MAY-1997 (TREMBLREL. 03, CREATED)
DT   01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT   01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE   ATP-DEPENDENT HELICASE.
OS   SULFOLOBUS SOLFATARICUS.
OC   ARCHAEBACTERIA; CRENARCHAEOTA; SULFOLOBALES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=P2;
RA   SENSEN C.W., KLENK H.P., SINGH R.K., ALLARD G., CHAN C.C.Y.,
RA   LIU Q.Y., PENNY S.L., YOUNG F., SCHENK M.E., GAASTERLAND T.,
RA   DOOLITTLE W.F., RAGAN M.A., CHARLEBOIS R.L.;
RL   MOL. MICROBIOL. 22:175-191(1996).
DR   EMBL; Y08257; E283825; -.
KW   HELICASE.
SQ   SEQUENCE   875 AA;  100392 MW;  5508D4B1 CRC32;

  Query Match           67.1%;  Score 55;  DB 9;  Length 875;
  Best Local Similarity  77.8%;  Pred. No. 1.09e+00;
  Matches    7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    193 LFRNLITKK 201
          ||| :|||| 
Qy      1 LFRAVITKK 9


RESULT    3
ID   Q91889         PRELIMINARY;   PRT;   693 AA.
AC   Q91889;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   UBIQUITIN-LIKE FUSION PROTEIN.
OS   XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93292985.
RA   LINNEN J.M., BAILEY C.P., WEEKS D.L.;
RL   GENE 128:181-188(1993).
DR   EMBL; L08474; G214866; -.
SQ   SEQUENCE   693 AA;  76862 MW;  3C04F5C3 CRC32;

  Query Match           65.9%;  Score 54;  DB 12;  Length 693;
  Best Local Similarity  50.0%;  Pred. No. 1.82e+00;
  Matches    6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    474 LFRSVEVRNIAD 485
          |||:|  :::||
Qy      1 LFRAVITKKVAD 12


RESULT    4
ID   Q91890         PRELIMINARY;   PRT;   701 AA.
AC   Q91890;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   UBIQUITIN-LIKE FUSION PROTEIN.
OS   XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93292985.
RA   LINNEN J.M., BAILEY C.P., WEEKS D.L.;
RL   GENE 128:181-188(1993).
DR   EMBL; L08475; G214868; -.
SQ   SEQUENCE   701 AA;  78581 MW;  369C8F29 CRC32;

  Query Match           65.9%;  Score 54;  DB 12;  Length 701;
  Best Local Similarity  50.0%;  Pred. No. 1.82e+00;
  Matches    6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    481 LFRSVEVRNIAD 492
          |||:|  :::||
Qy      1 LFRAVITKKVAD 12


RESULT    5
ID   Q22897         PRELIMINARY;   PRT;   362 AA.
AC   Q22897;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   COSMID C16D9.
GN   C16D9.3.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
```

```
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,
RA   COULSON A., CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A.,
RA   FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
RA   JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
RA   LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B.,
RA   O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A.,
RA   SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A., SONNHAMMER E.,
RA   STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M.,
RA   VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA   WILKINSON-SPROAT J., WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   GATTUNG S, LE TT.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   WATERSTON R.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U64858; G1465843; -.
SQ   SEQUENCE   362 AA;  42771 MW;  33408B5C CRC32;

  Query Match           64.6%;  Score 53;  DB 3;  Length 362;
  Best Local Similarity  33.3%;  Pred. No. 3.00e+00;
  Matches     4;  Conservative     6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    223 LFKEIASKRISD 234
          ||: : :|:::|
Qy      1 LFRAVITKKVAD 12


RESULT    6
ID   Q14798         PRELIMINARY;   PRT;   317 AA.
AC   Q14798;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   MAGE-4 PROTEIN.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95369706.
RA   IMAI Y., SHICHIJO S., YAMADA A., KATAYAMA T., YANO H., ITOH K.;
RL   GENE 160:287-290(1995).
DR   EMBL; D32075; G1125014; -.
SQ   SEQUENCE   317 AA;  35044 MW;  982B1AC9 CRC32;

  Query Match           63.4%;  Score 52;  DB 2;  Length 317;
  Best Local Similarity  41.7%;  Pred. No. 4.92e+00;
  Matches     5;  Conservative     4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    105 LFREALSNKVDE 116
          |||  :::|| :
Qy      1 LFRAVITKKVAD 12


RESULT    7
ID   O16057         PRELIMINARY;   PRT;   663 AA.
AC   O16057;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   LIM DOMAIN PROTEIN.
GN   ANON2D9.
OS   DROSOPHILA YAKUBA (FRUIT FLY).
OC   EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97420753.
RA   SCHMID K.J., TAUTZ D.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 94:9746-9750(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   SCHMID K.J., TAUTZ D.;
RL   SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF005862; G2459922; -.
SQ   SEQUENCE   663 AA;  76060 MW;  8F370705 CRC32;

  Query Match           63.4%;  Score 52;  DB 3;  Length 663;
  Best Local Similarity  50.0%;  Pred. No. 4.92e+00;
  Matches     6;  Conservative     3;  Mismatches   3;  Indels   0;  Gaps   0;

Db     65 LFRSTITEKLEN 76
          |||: || |: :
Qy      1 LFRAVITKKVAD 12


RESULT    8
ID   Q20049         PRELIMINARY;   PRT;   380 AA.
AC   Q20049;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   F35G12.2.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   CHUI C.;
RL   SUBMITTED (OCT-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,
RA   COULSON A., CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A.,
RA   FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
RA   JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
RA   LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B.,
RA   O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A.,
RA   SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A., SONNHAMMER E.,
RA   STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M.,
RA   VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA   WILKINSON-SPROAT J., WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
DR   EMBL; Z46242; G559425; -.
SQ   SEQUENCE   380 AA;  41659 MW;  BEADFCFB CRC32;

  Query Match           62.2%;  Score 51;  DB 3;  Length 380;
  Best Local Similarity  40.0%;  Pred. No. 8.02e+00;
  Matches     4;  Conservative     5;  Mismatches   1;  Indels   0;  Gaps   0;

Db    354 LFRTLVDKRI 363
          |||::: |::
Qy      1 LFRAVITKKV 10


RESULT    9
ID   O23469         PRELIMINARY;   PRT;   201 AA.
AC   O23469;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   SIMILARITY TO MYROSINASE-ASSOCIATED PROTEIN - BRASSICA.
OS   ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC   CAPPARALES; CRUCIFERAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   BEVAN M., STIEKEMA W., MURPHY G., WAMBUTT R., POHL T., TERRYN N.,
RA   KREIS M., KAVANAGH T., ENTIAN K.D., RIEGER M., JAMES R.,
RA   PUIGDOMENECH P., HATZOPOULOS P., OBERMAIER B., DUESTERHOFT A., JONES J.,
```

```
RA   PALME K., ANSORGE W., DELSENY M., BANCROFT I., MEWES H.W., SCHUELLER C.,
RA   CHALWATZIS N.;
RL   SUBMITTED (JUL-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU ARABIDOPSIS SEQUENCING PROJECT, ESSA;
RL   SUBMITTED (JUN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z97340; E326980; -.
SQ   SEQUENCE   201 AA;  22098 MW;  5F30ABCE CRC32;

  Query Match           59.8%;  Score 49;  DB 8;  Length 201;
  Best Local Similarity  45.5%;  Pred. No. 2.08e+01;
  Matches    5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    125 LYRQILAKEVA 135
          |:| :::| ||
Qy      1 LFRAVITKKVA 11


RESULT   10
ID   Q69112         PRELIMINARY;   PRT;   224 AA.
AC   Q69112;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   ULLORF.
GN   ULLORF.
OS   HERPES SIMPLEX VIRUS.
OC   VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89037371.
RA   WORRAD D.M., CARADONNA S.J.;
RL   J. VIROL. 62:4774-4777(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   WORRAD D.M., CARADONNA S.;
RL   SUBMITTED (FEB-1992) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   WORRAD D.M.;
RL   SUBMITTED (FEB-1992) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; M87011; G330304; -.
SQ   SEQUENCE   224 AA;  25128 MW;  BBCA07B0 CRC32;

  Query Match           59.8%;  Score 49;  DB 11;  Length 224;
  Best Local Similarity  50.0%;  Pred. No. 2.08e+01;
  Matches    6;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db     27 VLRSVIAKEVGD 38
          ::|:||:| |:|
Qy      1 LFRAVITKKVAD 12


RESULT   11
ID   Q58415         PRELIMINARY;   PRT;   236 AA.
AC   Q58415;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL PROTEIN MJ1009.
GN   MJ1009.
OS   METHANOCOCCUS JANNASCHII.
OC   ARCHAEBACTERIA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96337999.
RA   BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA   SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA   KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA   OVERBEEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA   SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA   UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA   COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA   KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RL   SCIENCE 273:1058-1073(1996).
CC   -!- SIMILARITY: TO E.COLI PHOSPHATE TRANSPORT SYSTEM REGULATORY
CC       PROTEIN.
DR   EMBL; U67543; G1591668; -.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   236 AA;  27426 MW;  0D3453CA CRC32;

  Query Match           59.8%;  Score 49;  DB 9;  Length 236;
  Best Local Similarity  33.3%;  Pred. No. 2.08e+01;
  Matches    4;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    161 LYRSMISKIIEN 172
          |:|::|:| : :
Qy      1 LFRAVITKKVAD 12


RESULT   12
ID   O24963         PRELIMINARY;   PRT;   287 AA.
AC   O24963;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 32.9 KD PROTEIN.
GN   HP0152.
OS   HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA;
OC   AEROBIC, MOTILE, HELICAL AND/OR VIBRIOID.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695;
RA   TOMB, WHITE, KERLAVAGE, CLAYTON, SUTTON, FLEISCHMANN, KETCHUM, KLENK,
RA   GILL, DOUGHERTY, NELSON, QUACKENBUSH, ZHOU, KIRKNESS, PETERSON, LOFTUS,
RA   RICHARDSON, DODSON, KHALAK, GLODEK, MCKENNEY, FITZEGERALD, LEE, ADAMS,
RA   HICKEY, BERG, GOCAYNE, UTTERBACK, PETERSON, KELLEY, COTTON, WEIDMAN,
RA   FUJII, BOWMAN, WATTHEY, WALLIN, HAYES, BORODOVSKY, KARP, SMITH,
RA   FRASER VENTER.;
RL   NATURE 388:539-547(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695;
RA   TOMB, WHITE, KERLAVAGE, CLAYTON, SUTTON, FLEISCHMANN, KETCHUM, KLENK,
RA   GILL, DOUGHERTY, NELSON, QUACKENBUSH, ZHOU, KIRKNESS, PETERSON, LOFTUS,
RA   RICHARDSON, DODSON, KHALAK, GLODEK, MCKENNEY, FITZEGERALD, LEE, ADAMS,
RA   HICKEY, BERG, GOCAYNE, UTTERBACK, PETERSON, KELLEY, COTTON, WEIDMAN,
RA   FUJII, BOWMAN, WATTHEY, WALLIN, HAYES, BORODOVSKY, KARP, SMITH,
RA   FRASER VENTER.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AE000536; G2313242; -.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   287 AA;  32948 MW;  81AA2EBE CRC32;

  Query Match           59.8%;  Score 49;  DB 9;  Length 287;
  Best Local Similarity  55.6%;  Pred. No. 2.08e+01;
  Matches    5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    189 LYRAILIKK 197
          |:||:: ||
Qy      1 LFRAVITKK 9


RESULT   13
ID   P91320         PRELIMINARY;   PRT;   446 AA.
AC   P91320;
DT   01-MAY-1997 (TREMBLREL. 03, CREATED)
DT   01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT   01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE   SIMILAR TO ACETYLCHOLINE RECEPTOR PROTEIN.
GN   F53E10.2.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   BECK C., WAMSLEY P.;
RL   SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   WATERSTON R.;
RL   SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U88177; G1825676; -.
SQ   SEQUENCE   446 AA;  52054 MW;  3317B8AC CRC32;

  Query Match           59.8%;  Score 49;  DB 3;  Length 446;
  Best Local Similarity  85.7%;  Pred. No. 2.08e+01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    140 LFRAIIT 146
          ||||:||
Qy      1 LFRAVIT 7


RESULT   14
ID   Q23501         PRELIMINARY;   PRT;   447 AA.
AC   Q23501;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   ZK455.8.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   WHITE S.;
RL   SUBMITTED (NOV-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
DR   EMBL; Z66567; G1051350; -.
SQ   SEQUENCE   447 AA;  50790 MW;  3DE8B3B0 CRC32;

  Query Match           59.8%;  Score 49;  DB 3;  Length 447;
  Best Local Similarity  66.7%;  Pred. No. 2.08e+01;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db    259 RSVITRKVG 267
          |:|||:||:
Qy      3 RAVITKKVA 11


RESULT   15
ID   Q19905         PRELIMINARY;   PRT;   481 AA.
AC   Q19905;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   F29F11.1.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   GARDNER A.;
RL   SUBMITTED (JUN-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
DR   EMBL; Z73974; E248315; -.
SQ   SEQUENCE   481 AA;  52755 MW;  4D9F8CBE CRC32;

  Query Match           59.8%;  Score 49;  DB 3;  Length 481;
  Best Local Similarity  54.5%;  Pred. No. 2.08e+01;
  Matches     6;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db    329 LFNTVTDKKIA 339
          || :|  ||:|
Qy      1 LFRAVITKKVA 11


RESULT   16
ID   Q92573         PRELIMINARY;   PRT;   522 AA.
AC   Q92573;
DT   01-FEB-1997 (TREMBLREL. 02, CREATED)
DT   01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   MYELOBLAST KIAA0241 (FRAGMENT).
GN   KIAA0241.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BONE MARROW;
RX   MEDLINE; 97191544.
RA   NAGASE T., SEKI N., ISHIKAWA K., OHIRA M., KAWARABAYASI Y., OHARA O.,
RA   TANAKA A., KOTANI H., MIYAJIMA N., NOMURA N.;
RL   DNA RES. 3:321-329(1996).
DR   EMBL; D87682; D1014125; -.
FT   NON_TER       1       1
SQ   SEQUENCE   522 AA;  57659 MW;  EC0B5136 CRC32;

  Query Match           59.8%;  Score 49;  DB 2;  Length 522;
  Best Local Similarity  46.2%;  Pred. No. 2.08e+01;
  Matches     6;  Conservative    5;  Mismatches   1;  Indels   1;  Gaps   1;

Db    456 LFRIVNVAKKIGN 468
          ||| | ::||:::
Qy      1 LFRAV-ITKKVAD 12


RESULT   17
```

```
ID   Q45923         PRELIMINARY;   PRT;   660 AA.
AC   Q45923;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   SIMILAR TO TNPB OF TRANSPOSON TN554 FROM STAPHYLOCOCCUS AUREUS.
OS   CLOSTRIDIUM BUTYRICUM.
OC   PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   HESSLINGER C., MULLER N., KAISER M., SAWERS G.;
RL   SUBMITTED (DEC-1993) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z29084; G436133; -.
SQ   SEQUENCE   660 AA;  78996 MW;  B9D4CCFC CRC32;

  Query Match           59.8%;  Score 49;  DB 9;  Length 660;
  Best Local Similarity  62.5%;  Pred. No. 2.08e+01;
  Matches     5;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db    210 FRVIVTKK 217
          || ::|||
Qy      2 FRAVITKK 9


RESULT   18
ID   Q63315         PRELIMINARY;   PRT;   813 AA.
AC   Q63315;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   LONG TYPE PB-CADHERIN.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=BRAIN;
RX   MEDLINE; 96212232.
RA   SUGIMOTO K., HONDA S., YAMAMOTO T., UEKI T., MONDEN M., KAJI A.,
RA   MATSUMOTO K., NAKAMURA T.;
RL   J. BIOL. CHEM. 271:11548-11556(1996).
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; D83348; G1398906; -.
DR   PROSITE; PS00232; CADHERIN; 2.
KW   CELL ADHESION; GLYCOPROTEIN; TRANSMEMBRANE; CALCIUM-BINDING; REPEAT.
SQ   SEQUENCE   813 AA;  87978 MW;  6E4D9F1D CRC32;

  Query Match           59.8%;  Score 49;  DB 10;  Length 813;
  Best Local Similarity  54.5%;  Pred. No. 2.08e+01;
  Matches     6;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    731 VFRDFISRKVA 741
          :||  |::|||
Qy      1 LFRAVITKKVA 11


RESULT   19
ID   O23524         PRELIMINARY;   PRT;   851 AA.
AC   O23524;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 96.8 KD PROTEIN.
OS   ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC   CAPPARALES; CRUCIFERAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   BEVAN M., STIEKEMA W., MURPHY G., WAMBUTT R., POHL T., TERRYN N.,
RA   KREIS M., KAVANAGH T., ENTIAN K.D., RIEGER M., JAMES R.,
RA   PUIGDOMENECH P., HATZOPOULOS P., OBERMAIER B., DUESTERHOFT A.,
RA   JONES J., PALME K., ANSORGE W., DELSENY M., BANCROFT I., MEWES H.W.,
RA   SCHUELLER C., CHALWATZIS N.;
RL   SUBMITTED (JUL-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU ARABIDOPSIS SEQUENCING PROJECT, ESSA;
RL   SUBMITTED (JUN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z97342; E327517; -.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   851 AA;  96849 MW;  23B0FD66 CRC32;

  Query Match           59.8%;  Score 49;  DB 8;  Length 851;
  Best Local Similarity  50.0%;  Pred. No. 2.08e+01;
  Matches     5;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    346 LFREMIDNKI 355
          ||| :| :|:
Qy      1 LFRAVITKKV 10


RESULT   20
ID   Q06164         PRELIMINARY;   PRT;  1454 AA.
AC   Q06164;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   CHROMOSOME XII COSMID 8543.
GN   L8543.13.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C (AB972);
RA   DU Z.;
RL   SUBMITTED (MAR-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C (AB972);
RA   JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,
RA   FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J.,
RA   KUCABA T., HALLSWORTH K., HAWKINS J., HILLIER L., JIER M.,
RA   JOHNSON D., JOHNSTON L., LANGSTON Y., LATREILLE P., LE T.,
RA   MARDIS E., MENEZES S., MILLER N., NHAN M., PAULEY A., PELUSO D.,
RA   RIFKEN L., RILES L., TAICH A., TREVASKIS E., VIGNATI D.,
RA   WILCOX L., WOHLDMAN P., VAUDIN M., WILSON R., WATERSTON R.;
RL   SUBMITTED (MAR-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C (AB972);
RA   WATERSTON R.;
RL   SUBMITTED (FEB-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U20618; G662137; -.
SQ   SEQUENCE   1454 AA;  167682 MW;  7C4BD4F7 CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 1454;
  Best Local Similarity  41.7%;  Pred. No. 2.08e+01;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    623 IFRVIIDKGITD 634
          :|| :| | ::|
Qy      1 LFRAVITKKVAD 12


RESULT   21
ID   Q61194         PRELIMINARY;   PRT;  1509 AA.
AC   Q61194;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE   P170 PHOSPHATIDYLINOSITOL 3-KINASE.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   VIRBASIUS J.V., GUILHERME A., CZECH M.P.;
RL   J. BIOL. CHEM. 271:13304-13307(1996).
DR   EMBL; U55772; G1305538; -.
SQ   SEQUENCE   1509 AA;  170777 MW;  E404092B CRC32;

  Query Match           59.8%;  Score 49;  DB 10;  Length 1509;
  Best Local Similarity  45.5%;  Pred. No. 2.08e+01;
  Matches     5;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db    545 LFKPIQSKKVG 555
          ||::: :|||:
Qy      1 LFRAVITKKVA 11


RESULT   22
ID   Q61182         PRELIMINARY;   PRT;  1658 AA.
AC   Q61182;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   PHOSPHOINOSITIDE 3-KINASE.
GN   CPK-M.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB C;
RA   MOLZ L.M., CHEN Y.W., HIRANO M., WILLIAMS L.T.;
RL   J. BIOL. CHEM. 271:13892-13899(1996).
DR   EMBL; U52193; G1272422; -.
SQ   SEQUENCE   1658 AA;  187439 MW;  D29F37A0 CRC32;

  Query Match           59.8%;  Score 49;  DB 10;  Length 1658;
  Best Local Similarity  45.5%;  Pred. No. 2.08e+01;
  Matches     5;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db    692 LFKPIQSKKVG 702
          ||::: :|||:
Qy      1 LFRAVITKKVA 11


RESULT   23
ID   O00443         PRELIMINARY;   PRT;  1686 AA.
AC   O00443;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   PHOSPHOINOSITIDE 3-KINASE.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   DOMIN J., PAGES F., VOLINIA S., RITTENHOUSE S.E., ZVELEBIL M.J.,
RA   STEIN R.C., WATERFIELD M.D.;
RL   BIOCHEM. J. 326:139-147(1997).
DR   EMBL; Y13367; E1188595; -.
SQ   SEQUENCE   1686 AA;  190736 MW;  8BA83E82 CRC32;

  Query Match           59.8%;  Score 49;  DB 2;  Length 1686;
  Best Local Similarity  45.5%;  Pred. No. 2.08e+01;
  Matches     5;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db    720 LFKPIQSKKVG 730
          ||::: :|||:
Qy      1 LFRAVITKKVA 11


RESULT   24
ID   O13661         PRELIMINARY;   PRT;  1807 AA.
AC   O13661;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 229.9KD PROTEIN IN NUC1-PRP21 INTERGENIC REGION.
GN   PI070.
OS   SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972 H-;
RA   KUSHIDA N., YAMAZAKI S., TANAKA T., JINNO K., HAIKAWA Y., YAMAZAKI J.,
RA   YAMAMOTO S., SEKINE M., OGUCHI A., NAGAI Y., SAKAI M., AOKI K.,
RA   OGURA K., OTSUKA R., KUDOH Y., YANAGIDA M., MACHIDA M., ZHANG M.Q.;
RL   SUBMITTED (JUL-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AB004539; D1022298; -.
SQ   SEQUENCE   1807 AA;  204620 MW;  F5887D03 CRC32;

  Query Match           59.8%;  Score 49;  DB 1;  Length 1807;
  Best Local Similarity  55.6%;  Pred. No. 2.08e+01;
  Matches     5;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    249 LFRDVIGRR 257
          ||| || ::
Qy      1 LFRAVITKK 9


RESULT   25
ID   Q98914         PRELIMINARY;   PRT;   131 AA.
AC   Q98914;
DT   01-FEB-1997 (TREMBLREL. 02, CREATED)
DT   01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT   01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
DE   CHICK OLFACTORY RECEPTOR 9 (FRAGMENT).
GN   COR9.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   NEF S., ALLAMAN I., FIUMELLI H., DE CASTRO E., NEF P.;
RL   MECH. DEV. 55:65-77(1996).
DR   EMBL; Z79592; E262005; -.
KW   G-PROTEIN COUPLED RECEPTOR.
FT   NON_TER       1      1
FT   NON_TER     131    131
SQ   SEQUENCE   131 AA;  14274 MW;  5FA99425 CRC32;

  Query Match           58.5%;  Score 48;  DB 12;  Length 131;
  Best Local Similarity  50.0%;  Pred. No. 3.31e+01;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db      8 LYSTVMTKRV 17
          |: :|:||:|
Qy      1 LFRAVITKKV 10


RESULT   26
ID   Q53037         PRELIMINARY;   PRT;   133 AA.
AC   Q53037;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   4.6KDA PROTEIN.
GN   NHHE.
OS   RHODOCOCCUS RHODOCHROUS.
OC   PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; NOCARDIOFORM.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J1;
RA   KOMEDA H., KOBAYASHI M., SHIMIZU S.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 93:4267-4272(1996).
DR   EMBL; D67027; G1402519; -.
SQ   SEQUENCE   133 AA;  14610 MW;  2AE0B4E6 CRC32;
```

```
  Query Match           58.5%;  Score 48;  DB 9;  Length 133;
  Best Local Similarity  41.7%;  Pred. No. 3.31e+01;
  Matches    5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db      3 VFRPVTTRRSSD 14
          :||:| |:: :|
Qy      1 LFRAVITKKVAD 12


RESULT   27
ID   P87594         PRELIMINARY;   PRT;   167 AA.
AC   P87594;
DT   01-MAY-1997 (TREMBLREL. 03, CREATED)
DT   01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT   01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE   P346, 54-KDA RNA DEPENDENT RNA POLYMERASE,
DE   P33, P6, P65, P61, P27, 25-KDA COAT PROTEIN (CPG), P18, P13, P20,
DE   AND P23 GENES, COMPLETE CDS.
OS   CITRUS TRISTEZA VIRUS (CTV).
OC   VIRIDAE; SS-RNA NONENVELOPED VIRUSES; CAPILLOVIRIDAE.
RN   [1]
RP   SEQUENCE OF 1-11 FROM N.A.
RC   STRAIN=VT;
RX   MEDLINE; 94106113.
RA   MAWASSI M., GAFNY R., BAR-JOSEPH M.;
RL   VIRUS GENES 7:265-275(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VT;
RX   MEDLINE; 95205086.
RA   MAWASSI M., GAFNY R., GAGLIARDI D., BAR-JOSEPH M.;
RL   J. GEN. VIROL. 76:651-659(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VT;
RA   MAWASSI M., MIETKIEWSKA E., GOFMAN R., YANG G., BAR-JOSEPH M.;
RL   J. GEN. VIROL. 77:2359-2364(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VT;
RA   BAR-JOSEPH M.;
RL   SUBMITTED (APR-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U56902; G1732502; -.
SQ   SEQUENCE   167 AA;  18266 MW;  CF2C717B CRC32;

  Query Match           58.5%;  Score 48;  DB 11;  Length 167;
  Best Local Similarity  41.7%;  Pred. No. 3.31e+01;
  Matches    5;  Conservative    3;  Mismatches   4;  Indels   0;  Gaps   0;

Db    124 LFKVIIGTNVDD 135
          ||: :|  :| |
Qy      1 LFRAVITKKVAD 12


RESULT   28
ID   Q66247         PRELIMINARY;   PRT;   167 AA.
AC   Q66247;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   P18, P13, P20, AND P23.5 GENES, COMPLETE CDS,
DE   AND COAT PROTEIN.
OS   CITRUS TRISTEZA VIRUS (CTV).
OC   VIRIDAE; SS-RNA NONENVELOPED VIRUSES; CAPILLOVIRIDAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VT;
RX   MEDLINE; 95205086.
RA   MAWASSI M., GAFNY R., GAGLIARDI D., BAR-JOSEPH M.;
RL   J. GEN. VIROL. 76:651-659(1995).
DR   EMBL; U17265; G596063; -.
SQ   SEQUENCE   167 AA;  18292 MW;  93BD9214 CRC32;

  Query Match           58.5%;  Score 48;  DB 11;  Length 167;
  Best Local Similarity  41.7%;  Pred. No. 3.31e+01;
  Matches    5;  Conservative    3;  Mismatches   4;  Indels   0;  Gaps   0;

Db    124 LFKVIIGTNVDD 135
          ||: :|  :| |
Qy      1 LFRAVITKKVAD 12


RESULT   29
ID   O21267         PRELIMINARY;   PRT;   281 AA.
AC   O21267;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   ATP SYNTHASE F1 SUBUNIT ALPHA.
GN   ATP3.
OS   RECLINOMONAS AMERICANA.
OG   MITOCHONDRION.
OC   EUKARYOTA; NOT YET CLASSIFIED.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC50394;
RX   MEDLINE; 97311393.
RA   LANG B.F., BURGER G., O'KELLY C.J., CEDERGREN R., GOLDING G.B.,
RA   LEMIEUX C., SANKOFF D., TURMEL M., GRAY M.W.;
RL   NATURE 387:493-497(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC50394;
RA   LANG B.F., BURGER G.;
RL   SUBMITTED (JUN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF007261; G2258360; -.
KW   MITOCHONDRION.
SQ   SEQUENCE   281 AA;  31195 MW;  87B481F4 CRC32;

  Query Match           58.5%;  Score 48;  DB 6;  Length 281;
  Best Local Similarity  45.5%;  Pred. No. 3.31e+01;
  Matches    5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    166 FRSVLTQNVIE 176
          ||:|:| :| :
Qy      2 FRAVITKKVAD 12


RESULT   30
ID   O28219         PRELIMINARY;   PRT;   319 AA.
AC   O28219;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   ACTIVATOR 1, REPLICATION FACTOR C, 35 KD SUBUNIT.
GN   AF2060.
OS   ARCHAEOGLOBUS FULGIDUS.
OC   ARCHAEBACTERIA; EURYARCHAEOTA; ARCHAEOGLOBALES; ARCHAEOGLOBACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   KLENK H.P., CLAYTON R.A., TOMB J., WHITE O., NELSON K.E.,
RA   KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,
RA   RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,
RA   FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
RA   KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
RA   PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,
RA   OVERBEEK R., GOCAYNE J.D., WEIDMAN J.F., MCDONALD L., UTTERBACK T.,
RA   COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,
RA   SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,
RA   MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,
RA   VENTER J.C.;
RL   SUBMITTED (DEC-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   KLENK H.P., CLAYTON R.A., TOMB J., WHITE O., NELSON K.E.,
RA   KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,
```

```
RA   RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,
RA   FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
RA   KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
RA   PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,
RA   OVERBEEK R., GOCAYNE J.D., WEIDMAN J.F., MCDONALD L., UTTERBACK T.,
RA   COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,
RA   SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,
RA   MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,
RA   VENTER J.C.;
RL   SUBMITTED (NOV-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AE000961; G2648471; -.
SQ   SEQUENCE   319 AA;  35993 MW;  BC82CD7C CRC32;

  Query Match           58.5%;  Score 48;  DB 9;  Length 319;
  Best Local Similarity  41.7%;  Pred. No. 3.31e+01;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    268 LFREIISMPIKD 279
          ||| :|:: : |
Qy      1 LFRAVITKKVAD 12


RESULT   31
ID   Q58293         PRELIMINARY;   PRT;   336 AA.
AC   Q58293;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL PROTEIN MJ0883.
GN   MJ0883.
OS   METHANOCOCCUS JANNASCHII.
OC   ARCHAEBACTERIA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96337999.
RA   BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA   SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA   KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA   OVERBEEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA   SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA   UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA   COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA   KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RL   SCIENCE 273:1058-1073(1996).
CC   -!- SIMILARITY: STRONG IN THE C-TERMINAL, OF M.JANNASCHII M1557.
DR   EMBL; U67532; G1499713; -.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; UNKNOWN_1.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   336 AA;  39000 MW;  7F372A03 CRC32;

  Query Match           58.5%;  Score 48;  DB 9;  Length 336;
  Best Local Similarity  62.5%;  Pred. No. 3.31e+01;
  Matches     5;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db     76 FREIISKK 83
          || :|:||
Qy      2 FRAVITKK 9


RESULT   32
ID   O26038         PRELIMINARY;   PRT;   458 AA.
AC   O26038;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   FERRODOXIN-LIKE PROTEIN.
GN   HP1508.
OS   HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA;
OC   AEROBIC, MOTILE, HELICAL AND/OR VIBRIOID.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695;
RA   TOMB, WHITE, KERLAVAGE, CLAYTON, SUTTON, FLEISCHMANN, KETCHUM, KLENK,
RA   GILL, DOUGHERTY, NELSON, QUACKENBUSH, ZHOU, KIRKNESS, PETERSON,
RA   LOFTUS, RICHARDSON, DODSON, KHALAK, GLODEK, MCKENNEY, FITZEGERALD,
RA   LEE, ADAMS, HICKEY, BERG, GOCAYNE, UTTERBACK, PETERSON, KELLEY,
RA   COTTON, WEIDMAN, FUJII, BOWMAN, WATTHEY, WALLIN, HAYES, BORODOVSKY,
RA   KARP, SMITH, FRASER VENTER.;
RL   NATURE 388:539-547(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695;
RA   TOMB, WHITE, KERLAVAGE, CLAYTON, SUTTON, FLEISCHMANN, KETCHUM, KLENK,
RA   GILL, DOUGHERTY, NELSON, QUACKENBUSH, ZHOU, KIRKNESS, PETERSON,
RA   LOFTUS, RICHARDSON, DODSON, KHALAK, GLODEK, MCKENNEY, FITZEGERALD,
RA   LEE, ADAMS, HICKEY, BERG, GOCAYNE, UTTERBACK, PETERSON, KELLEY,
RA   COTTON, WEIDMAN, FUJII, BOWMAN, WATTHEY, WALLIN, HAYES, BORODOVSKY,
RA   KARP, SMITH, FRASER VENTER.;
RL   SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AE000649; G2314689; -.
DR   PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW   IRON-SULFUR.
SQ   SEQUENCE   458 AA;  52651 MW;  80152970 CRC32;

  Query Match           58.5%;  Score 48;  DB 9;  Length 458;
  Best Local Similarity  50.0%;  Pred. No. 3.31e+01;
  Matches     5;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db    100 LYRDVIETKI 109
          |:| ||  |:
Qy      1 LFRAVITKKV 10


RESULT   33
ID   O32912         PRELIMINARY;   PRT;   478 AA.
AC   O32912;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE.
GN   GUAB.
OS   MYCOBACTERIUM LEPRAE.
OC   PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   SKELTON J., CHURCHER C.M.;
RL   SUBMITTED (OCT-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL   SUBMITTED (OCT-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   EIGLMEIER K., HONORE N., WOODS S.A., CAUDRON B., COLE S.T.;
RL   MOL. MICROBIOL. 7:197-206(1993).
DR   EMBL; AL008609; E1168632; -.
SQ   SEQUENCE   478 AA;  50383 MW;  B8FC1940 CRC32;

  Query Match           58.5%;  Score 48;  DB 9;  Length 478;
  Best Local Similarity  60.0%;  Pred. No. 3.31e+01;
  Matches     6;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    390 RAVVARTVAD 399
          |||::: |||
Qy      3 RAVITKKVAD 12


RESULT   34
ID   O04130         PRELIMINARY;   PRT;   624 AA.
AC   O04130;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)
DE   PHOSPHOGLYCERATE DEHYDROGENASE.
OS   ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
```

```
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC   CAPPARALES; CRUCIFERAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   HO C., NOJI M., SAITO K.;
RL   SUBMITTED (APR-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AB003280; D1021238; -.
SQ   SEQUENCE   624 AA;  66453 MW;  598FBAC6 CRC32;

  Query Match           58.5%;  Score 48;  DB 8;  Length 624;
  Best Local Similarity  50.0%;  Pred. No. 3.31e+01;
  Matches    6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    450 LLRAMITKGIIE 461
          |:||:||| : :
Qy      1 LFRAVITKKVAD 12


RESULT   35
ID   P73465         PRELIMINARY;   PRT;   793 AA.
AC   P73465;
DT   01-FEB-1997 (TREMBLREL. 02, CREATED)
DT   01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT   01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
DE   DNA HELICASE II.
GN   UVRD.
OS   SYNECHOCYSTIS SP.
OC   EUBACTERIA; CYANOBACTERIA; CHROOCOCCALES; SYNECHOCYSTIS.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PCC6803;
RA   TABATA S.;
RL   SUBMITTED (JUN-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PCC6803;
RA   KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA   MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA   HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA   SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA   TABATA S.;
RL   DNA RES. 3:109-136(1996).
DR   EMBL; D90906; G1652584; -.
KW   HELICASE.
SQ   SEQUENCE   793 AA;  90395 MW;  47FD4B3F CRC32;

  Query Match           58.5%;  Score 48;  DB 9;  Length 793;
  Best Local Similarity  40.0%;  Pred. No. 3.31e+01;
  Matches    4;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db    160 LFKTIVTKDM 169
          ||::::|| :
Qy      1 LFRAVITKKV 10


Search completed: Tue Apr  7 08:40:13 1998
Job time : 14 secs.
```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:42:04 1998; MasPar time 6.36 Seconds
79.499 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-4
Description: (1-12) from 5541104.pep
Perfect Score: 81
Sequence: 1 DVKEADPTGHSY 12

Scoring table: PAM 150
Gap 15

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: sptrembl5
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phage 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 22.764; Variance 22.337; scale 1.019

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	63	77.8	347	2	O00601	DAM10-DSS-AHC CRITICAL	1.78e-04
2	59	72.8	317	2	Q14798	MAGE-4 PROTEIN.	2.52e-03
3	58	71.6	330	10	Q60762	MELANOMA ANTIGEN, RELA	4.83e-03
4	58	71.6	330	10	Q60763	MELANOMA ANTIGEN, RELA	4.83e-03
5	58	71.6	380	10	Q60761	MELANOMA ANTIGEN, RELA	4.83e-03
6	54	66.7	346	2	O15481	MAGE-B4.	6.07e-02
7	53	65.4	277	3	Q17687	C06A6.3 PROTEIN.	1.12e-01
8	51	63.0	334	9	O29369	SIGNAL-TRANSDUCING HIS	3.77e-01
9	48	59.3	90	4	Q28835	CGMP-GATED RETINAL PHO	2.18e+00
10	48	59.3	308	2	Q92788	RAD GTPASE.	2.18e+00
11	48	59.3	309	9	Q44483	NIFU2.	2.18e+00
12	48	59.3	450	9	O07752	HYPOTHETICAL 46.7 KD P	2.18e+00
13	48	59.3	875	9	O31978	YOMG PROTEIN.	2.18e+00
14	48	59.3	4340	9	O30764	POLYKETIDE SYNTHASE MO	2.18e+00
15	47	58.0	215	1	Q09717	RAN/SPI1 BINDING PROTE	3.84e+00
16	47	58.0	313	2	O15479	MAGE-B2.	3.84e+00
17	46	56.8	108	9	P74837	NAPHTHALENESULFONIC AC	6.72e+00
18	46	56.8	198	9	P95779	DTDP-4-KETO-L-RHAMNOSE	6.72e+00
19	46	56.8	272	3	O17311	20S PROTEASOME BETA2 S	6.72e+00
20	46	56.8	272	3	O17312	20S PROTEASOME BETA2 S	6.72e+00
21	46	56.8	317	8	Q41681	1-AMINOCYCLOPROPANE-1-	6.72e+00
22	46	56.8	356	8	P92966	SPLICING FACTOR.	6.72e+00
23	46	56.8	382	9	Q51532	PILU.	6.72e+00
24	46	56.8	439	8	Q40565	RIBULOSE BISPHOSPHATE	6.72e+00
25	46	56.8	526	3	Q24044	FIZZY (FZY).	6.72e+00
26	46	56.8	543	1	Q05664	TRANSCRIPTION FACTOR.	6.72e+00
27	46	56.8	551	1	Q08224	CHROMOSOME XV READING	6.72e+00
28	46	56.8	551	1	Q08975	ORF YPL258C.	6.72e+00
29	46	56.8	597	10	Q63615	VACUOLAR PROTEIN SORTI	6.72e+00
30	46	56.8	628	7	Q38231	UNIDENTIFIED ORF34.	6.72e+00
31	46	56.8	702	3	Q93539	F20D1.7.	6.72e+00
32	46	56.8	999	2	Q12866	CELLULAR PROTO-ONCOGEN	6.72e+00
33	46	56.8	1187	9	Q59278	ENDOXYLANASE (EC 3.2.1	6.72e+00
34	45	55.6	46	9	O34568	YBBU PROTEIN.	1.16e+01
35	45	55.6	86	10	Q61740	INTEGRIN ALPHA 4 (CD49	1.16e+01
36	45	55.6	125	9	Q46451	RIBOSOMAL PROTEIN S11.	1.16e+01
37	45	55.6	199	9	O32481	GRPE.	1.16e+01
38	45	55.6	353	11	Q88626	ORFIV.	1.16e+01
39	45	55.6	477	9	O32354	GLUTAMINE SYNTHETASE (	1.16e+01
40	45	55.6	503	11	O10400	P57 (FRAGMENT).	1.16e+01
41	45	55.6	503	11	O10399	P57 (FRAGMENT).	1.16e+01
42	45	55.6	503	11	O10397	P57 (FRAGMENT).	1.16e+01
43	45	55.6	503	11	O10403	P57 (FRAGMENT).	1.16e+01
44	45	55.6	503	11	O10394	P57 (FRAGMENT).	1.16e+01
45	45	55.6	503	11	Q89857	GLYCOPROTEIN.	1.16e+01
46	45	55.6	503	11	Q65459	GLYCOPROTEIN.	1.16e+01
47	45	55.6	581	9	Q51317	PENICILLIN-BINDING PRO	1.16e+01
48	45	55.6	582	9	Q50960	PENICILLIN-BINDING PRO	1.16e+01
49	45	55.6	644	9	P77391	FROM BASES 1860594 TO	1.16e+01
50	45	55.6	770	9	O33622	PBP PROTEIN.	1.16e+01
51	45	55.6	1032	10	Q61989	INTEGRIN ALPHA 4 (CD49	1.16e+01
52	45	55.6	1078	10	O08995	MYELIN TRANSCRIPTION F	1.16e+01
53	45	55.6	1288	2	O15038	KIAA0324 (FRAGMENT).	1.16e+01
54	45	55.6	1518	3	Q21850	SIMILAR TO LDL RECEPTO	1.16e+01
55	44	54.3	61	9	Q54538	M TYPE 15 (EMML) (FRAG	1.99e+01
56	44	54.3	221	12	Q90685	AGRIN (FRAGMENT).	1.99e+01
57	44	54.3	377	9	Q53475	EMML15.	1.99e+01
58	44	54.3	438	11	Q66640	VIRION PROTEIN KINASE.	1.99e+01
59	44	54.3	449	9	P95692	GLUTAMINE SYNTHETASE T	1.99e+01
60	44	54.3	467	9	P73043	GLUTAMATE DECARBOXYLAS	1.99e+01
61	44	54.3	608	11	O41253	PNS7.	1.99e+01
62	44	54.3	608	9	O28171	SIGNAL-TRANSDUCING HIS	1.99e+01
63	44	54.3	775	11	P89057	RNA FOR VP4, COMPLETE	1.99e+01
64	44	54.3	2044	3	P91620	STILL LIFE TYPE 2.	1.99e+01
65	44	54.3	2457	11	O41965	TEGUMENT PROTEIN.	1.99e+01
66	43	53.1	67	3	Q26797	PROTEIN KINASE (FRAGME	3.38e+01
67	43	53.1	156	3	Q23961	AMYLASE (FRAGMENT).	3.38e+01
68	43	53.1	272	10	O35830	LECITHIN-CHOLESTEROL A	3.38e+01
69	43	53.1	330	10	P97269	SPERAD-B.	3.38e+01
70	43	53.1	337	10	P97268	SPERAD-A.	3.38e+01
71	43	53.1	341	3	Q25028	CYSTEINE PROTEINASE.	3.38e+01
72	43	53.1	370	12	Q91904	FORK HEAD PROTEIN.	3.38e+01
73	43	53.1	390	9	O28032	ARGININOSUCCINATE SYNT	3.38e+01
74	43	53.1	506	3	Q17381	CE-FKH-1.	3.38e+01
75	43	53.1	564	2	Q14111	MATRIX METALLOPROTEINA	3.38e+01
76	43	53.1	615	7	O21976	GP49.	3.38e+01
77	43	53.1	621	11	Q66196	PEPLOMERIC POLYPROTEIN	3.38e+01
78	43	53.1	762	10	Q64595	CGMP DEPENDENT PROTEIN	3.38e+01
79	43	53.1	767	3	O15882	POLO-LIKE PROTEIN KINA	3.38e+01
80	43	53.1	2509	2	Q16702	FATTY ACID SYNTHASE (E	3.38e+01
81	43	53.1	3224	2	Q15280	RANBP2 (EC 5.2.1.8) (R	3.38e+01
82	42	51.9	156	11	Q84828	17 KD PROTEIN.	5.68e+01
83	42	51.9	156	11	Q84839	ORF 5.	5.68e+01
84	42	51.9	288	9	Q51323	MOLYBDENUM-CONTAINING	5.68e+01
85	42	51.9	307	9	O24828	CATECHOL 2,3-DIOXYGENA	5.68e+01
86	42	51.9	349	9	Q59575	TUNGSTEN FORMYLMETHANO	5.68e+01
87	42	51.9	381	3	Q94399	ZK265.2.	5.68e+01
88	42	51.9	401	13	O36039	ELONGATION FACTOR-1 AL	5.68e+01
89	42	51.9	444	3	Q21861	R09D1.5.	5.68e+01
90	42	51.9	445	9	P94491	XYLOSE ISOMERASE (EC 5	5.68e+01
91	42	51.9	446	8	O23175	HYPOTHETICAL 50.3 KD P	5.68e+01
92	42	51.9	461	9	P71984	HYPOTHETICAL 49.4 KD P	5.68e+01
93	42	51.9	493	2	Q92772	P56 KKIAMRE PROTEIN KI	5.68e+01

94	42	51.9	503	9	O33360	HYPOTHETICAL 54.2 KD P	5.68e+01
95	42	51.9	627	11	Q65811	(STRAIN DRAPER) P125 (	5.68e+01
96	42	51.9	691	3	O01477	CODED FOR BY C. ELEGAN	5.68e+01
97	42	51.9	859	3	Q21401	COSMID K09E3.	5.68e+01
98	42	51.9	1358	11	Q65814	(STRAIN SINGER) P125 (	5.68e+01
99	42	51.9	1358	11	Q65813	(STRAIN OREGON) P125 (	5.68e+01
100	42	51.9	1608	7	O03937	MINOR CAPSID PROTEIN.	5.68e+01

ALIGNMENTS

```
RESULT    1
ID   O00601         PRELIMINARY;   PRT;   347 AA.
AC   O00601;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)
DE   DAM10=DSS-AHC CRITICAL INTERVAL MAGE SUPERFAMILY GENE 10.
GN   DAM10.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96081328.
RA   DABOVIC B., ZANARIA E., BARDONI B., LISA A., BORDIGNON C., RUSSO V.,
RA   MATESSI C., TRAVERSARI C., CAMERINO G.;
RL   MAMM. GENOME 6:571-580(1995).
DR   EMBL; S80936; E323784; -.
SQ   SEQUENCE   347 AA;  39049 MW;  AB96D5BB CRC32;

  Query Match           77.8%;  Score 63;  DB 2;  Length 347;
  Best Local Similarity  58.3%;  Pred. No. 1.78e-04;
  Matches     7;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    164 DLKEDNPSGHTY 175
          |:|| :|:||:|
Qy      1 DVKEADPTGHSY 12


RESULT    2
ID   Q14798         PRELIMINARY;   PRT;   317 AA.
AC   Q14798;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   MAGE-4 PROTEIN.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95369706.
RA   IMAI Y., SHICHIJO S., YAMADA A., KATAYAMA T., YANO H., ITOH K.;
RL   GENE 160:287-290(1995).
DR   EMBL; D32075; G1125014; -.
SQ   SEQUENCE   317 AA;  35044 MW;  982B1AC9 CRC32;

  Query Match           72.8%;  Score 59;  DB 2;  Length 317;
  Best Local Similarity  58.3%;  Pred. No. 2.52e-03;
  Matches     7;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    166 DVKEVDPASNTY 177
          |||| ||:::::|
Qy      1 DVKEADPTGHSY 12


RESULT    3
ID   Q60762         PRELIMINARY;   PRT;   330 AA.
AC   Q60762;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   MELANOMA ANTIGEN, RELATED SEQUENCE 2 (SMAGE-2 PROTEIN).
GN   MAGE-RS2.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2; TISSUE=KIDNEY;
RX   MEDLINE; 96070435.
RA   DE BACKER O., VERHEYDEN A.M., MARTIN B., GODELAINE D., DE PLAEN E.,
RA   BRASSEUR R., AVNER P., BOON T.;
RL   GENOMICS 28:74-83(1995).
DR   EMBL; U19032; G1165172; -.
DR   MGD; MGI:105117; MAGE-RS2.
SQ   SEQUENCE   330 AA;  35936 MW;  36D760C5 CRC32;

  Query Match           71.6%;  Score 58;  DB 10;  Length 330;
  Best Local Similarity  58.3%;  Pred. No. 4.83e-03;
  Matches     7;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    158 ELKEIDPSTHSY 169
          ::|| ||: |||
Qy      1 DVKEADPTGHSY 12


RESULT    4
ID   Q60763         PRELIMINARY;   PRT;   330 AA.
AC   Q60763;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   MELANOMA ANTIGEN, RELATED SEQUENCE 2 (SMAGE-3 PROTEIN).
GN   MAGE-RS2.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2; TISSUE=KIDNEY;
RX   MEDLINE; 96070435.
RA   DE BACKER O., VERHEYDEN A.M., MARTIN B., GODELAINE D.,
RA   DE PLAEN E., BRASSEUR R., AVNER P., BOON T.;
RL   GENOMICS 28:74-83(1995).
DR   EMBL; U19033; G1165174; -.
DR   MGD; MGI:105117; MAGE-RS2.
SQ   SEQUENCE   330 AA;  35985 MW;  83AD4246 CRC32;

  Query Match           71.6%;  Score 58;  DB 10;  Length 330;
  Best Local Similarity  58.3%;  Pred. No. 4.83e-03;
  Matches     7;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    158 ELKEIDPSTHSY 169
          ::|| ||: |||
Qy      1 DVKEADPTGHSY 12


RESULT    5
ID   Q60761         PRELIMINARY;   PRT;   380 AA.
AC   Q60761;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   MELANOMA ANTIGEN, RELATED SEQUENCE 2 (SMAGE-1 PROTEIN).
GN   MAGE-RS2.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2; TISSUE=KIDNEY;
RX   MEDLINE; 96070435.
RA   DE BACKER O., VERHEYDEN A.M., MARTIN B., GODELAINE D., DE PLAEN E.,
RA   BRASSEUR R., AVNER P., BOON T.;
```

```
RL   GENOMICS 28:74-83(1995).
DR   EMBL; U19031; G1165170; -.
DR   MGD; MGI:105117; MAGE-RS2.
SQ   SEQUENCE   380 AA;  41317 MW;  AD78F38E CRC32;

  Query Match           71.6%;  Score 58;  DB 10;  Length 380;
  Best Local Similarity  58.3%;  Pred. No. 4.83e-03;
  Matches     7;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    208 ELKEIDPSTHSY 219
          ::|| ||: |||
Qy      1 DVKEADPTGHSY 12


RESULT    6
ID   O15481         PRELIMINARY;   PRT;   346 AA.
AC   O15481;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   MAGE-B4.
GN   MAGE-B4.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   LURQUIN C.;
RL   SUBMITTED (OCT-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U93163; G2459682; -.
SQ   SEQUENCE   346 AA;  38923 MW;  C6A5A407 CRC32;

  Query Match           66.7%;  Score 54;  DB 2;  Length 346;
  Best Local Similarity  63.6%;  Pred. No. 6.07e-02;
  Matches     7;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    166 LKEVNPTTHSY 176
          :|| :|| |||
Qy      2 VKEADPTGHSY 12


RESULT    7
ID   Q17687         PRELIMINARY;   PRT;   277 AA.
AC   Q17687;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   C06A6.3 PROTEIN.
GN   C06A6.3.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   WU X., GATTUNG S.;
RL   SUBMITTED (DEC-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U41012; G1086626; -.
SQ   SEQUENCE   277 AA;  30379 MW;  D0A604DF CRC32;

  Query Match           65.4%;  Score 53;  DB 3;  Length 277;
  Best Local Similarity  54.5%;  Pred. No. 1.12e-01;
  Matches     6;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db     87 VRESDPIPHGY 97
          |:|:||  |:|
Qy      2 VKEADPTGHSY 12


RESULT    8
ID   O29369         PRELIMINARY;   PRT;   334 AA.
AC   O29369;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   SIGNAL-TRANSDUCING HISTIDINE KINASE.
GN   AF0893.
OS   ARCHAEOGLOBUS FULGIDUS.
OC   ARCHAEBACTERIA; EURYARCHAEOTA; ARCHAEOGLOBALES; ARCHAEOGLOBACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   KLENK H.P., CLAYTON R.A., TOMB J., WHITE O., NELSON K.E., KETCHUM K.A.,
RA   DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D., RICHARDSON D.L.,
RA   KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C., FLEISCHMANN R.D.,
RA   QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S., KIRKNESS E.F.,
RA   DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B., PETERSON S.,
RA   REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L., OVERBEEK R.,
RA   GOCAYNE J.D., WEIDMAN J.F., MCDONALD L., UTTERBACK T., COTTON M.D.,
RA   SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M., SADOW P.W.,
RA   D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A., MASON T.M.,
RA   OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RL   SUBMITTED (DEC-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   KLENK H.P., CLAYTON R.A., TOMB J., WHITE O., NELSON K.E., KETCHUM K.A.,
RA   DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D., RICHARDSON D.L.,
RA   KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C., FLEISCHMANN R.D.,
RA   QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S., KIRKNESS E.F.,
RA   DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B., PETERSON S.,
RA   REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L., OVERBEEK R.,
RA   GOCAYNE J.D., WEIDMAN J.F., MCDONALD L., UTTERBACK T., COTTON M.D.,
RA   SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M., SADOW P.W.,
RA   D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A., MASON T.M.,
RA   OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RL   SUBMITTED (NOV-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AE001042; G2649708; -.
SQ   SEQUENCE   334 AA;  38033 MW;  9E77911F CRC32;

  Query Match           63.0%;  Score 51;  DB 9;  Length 334;
  Best Local Similarity  41.7%;  Pred. No. 3.77e-01;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    314 EVKDNEPTGTVF 325
          :||: :|||  :
Qy      1 DVKEADPTGHSY 12


RESULT    9
ID   Q28835         PRELIMINARY;   PRT;    90 AA.
AC   Q28835;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   CGMP-GATED RETINAL PHOTORECEPTOR CHANNEL (FRAGMENT).
OS   ORYCTOLAGUS CUNICULUS (RABBIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; LAGOMORPHA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   HUNDAL S.P., DIFRANCESCO D., MANGONI M., BRAMMAR W.J., CONLEY E.C.;
RL   BIOCHEM. SOC. TRANS. 21:0-0(1993).
DR   EMBL; S65218; G410552; -.
FT   NON_TER       1      1
SQ   SEQUENCE   90 AA;  10981 MW;  98DBC53A CRC32;

  Query Match           59.3%;  Score 48;  DB 4;  Length 90;
  Best Local Similarity  41.7%;  Pred. No. 2.18e+00;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db     22 EVMIIDPSGNTY 33
          :|:  ||:|::|
Qy      1 DVKEADPTGHSY 12


RESULT   10
ID   Q92788         PRELIMINARY;   PRT;   308 AA.
AC   Q92788;
```

```
DT   01-FEB-1997 (TREMBLREL. 02, CREATED)
DT   01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   RAD GTPASE.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94069319.
RA   REYNET C., KAHN C.R.;
RL   SCIENCE 262:1441-1444(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96375161.
RA   CALDWELL J.S., MOYERS J.S., DORIA A., REYNET C., KAHN R.C.;
RL   BIOCHIM. BIOPHYS. ACTA 1316:145-148(1996).
DR   EMBL; U46165; G1620563; -.
SQ   SEQUENCE   308 AA;  33220 MW;  4B7A3673 CRC32;

  Query Match           59.3%;  Score 48;  DB 2;  Length 308;
  Best Local Similarity  55.6%;  Pred. No. 2.18e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    119 EAEAAGHTY 127
          ||:::||:|
Qy      4 EADPTGHSY 12


RESULT   11
ID   Q44483         PRELIMINARY;   PRT;   309 AA.
AC   Q44483;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   NIFU2.
GN   NIFU2.
OS   ANABAENA VARIABILIS.
OC   PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC   CYANOBACTERIA (BLUE-GREEN ALGAE); NOSTOCALES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29413;
RX   MEDLINE; 96016168.
RA   THIEL T., LYONS E.M., ERKER J.C., ERNST A.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 92:9358-9362(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29413;
RA   THIEL T., LYONS E.M., ERKER J.C.;
RL   SUBMITTED (FEB-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U49859; G1236928; -.
SQ   SEQUENCE   309 AA;  33028 MW;  4899DDAA CRC32;

  Query Match           59.3%;  Score 48;  DB 9;  Length 309;
  Best Local Similarity  54.5%;  Pred. No. 2.18e+00;
  Matches     6;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db    194 IKESAPVGTSY 204
          :||: | | ||
Qy      2 VKEADPTGHSY 12


RESULT   12
ID   O07752         PRELIMINARY;   PRT;   450 AA.
AC   O07752;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 46.7 KD PROTEIN.
GN   MTCY180.40C.
OS   MYCOBACTERIUM TUBERCULOSIS.
OC   PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RA   OLIVER K., HARRIS D.;
RL   SUBMITTED (JUN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RA   BARRELL B.G., RAJANDREAM M.A., PARKHILL J.;
RL   SUBMITTED (JUN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE; 96181548.
RA   PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
RA   BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA   COLE S.T.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).
DR   EMBL; Z97193; E324883; -.
KW   HYPOTHETICAL PROTEIN.
SQ   SEQUENCE   450 AA;  46720 MW;  1606486B CRC32;

  Query Match           59.3%;  Score 48;  DB 9;  Length 450;
  Best Local Similarity  41.7%;  Pred. No. 2.18e+00;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db    337 EVKVVDPSANPY 348
          :||  ||::::|
Qy      1 DVKEADPTGHSY 12


RESULT   13
ID   O31978         PRELIMINARY;   PRT;   875 AA.
AC   O31978;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   YOMG PROTEIN.
GN   YOMG.
OS   BACILLUS SUBTILIS.
OC   PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
RA   AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
RA   BORRISS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
RA   BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
RA   CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
RA   DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,
RA   ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D., FRITZ C.,
RA   FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N., GHIM S.Y.,
RA   GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G., GUISEPPI G., GUY B.J.,
RA   HAGA K., HAIECH J., HARWOOD C.R., HENAUT A., HILBERT H., HOLSAPPEL S.,
RA   HOSONO S., HULLO M.F., ITAYA M., JONES L., JORIS B., KARAMATA D.,
RA   KASAHARA Y., KLAERR-BLANCHARD M., KLEIN C., KOBAYASHI Y., KOETTER P.,
RA   KONINGSTEIN G., KROGH S., KUMANO M., KURITA K., LAPIDUS A.,
RA   LARDINOIS S., LAUBER J., LAZAREVIC V., LEE S.M., LEVINE A., LIU H.,
RA   MASUDA S., MAUEL C., MEDIGUE C., MEDINA N., MELLADO R.P., MIZUNO M.,
RA   MOESTL D., NAKAI S., NOBACK M., NOONE D., O'REILLY M., OGAWA K.,
RA   OGIWARA A., OUDEGA B., PARK S.H., PARRO V., POHL T.M., PORTETELLE D.,
RA   PORWOLLIK S., PRESCOTT A.M., PRESECAN E., PUJIC P., PURNELLE B.,
RA   RAPOPORT G., REY M., REYNOLDS S., RIEGER M., RIVOLTA C., ROCHA E.,
RA   ROCHE B., ROSE M., SADAIE Y., SATO T., SCANLAN E., SCHLEICH S.,
RA   SCHROETER R., SCOFFONE F., SEKIGUCHI J., SEKOWSKA A., SEROR S.J.,
RA   SERROR P., SHIN B.S., SOLDO B., SOROKIN A., TACCONI E., TAKAGI T.,
RA   TAKAHASHI H., TAKEMARU K., TAKEUCHI M., TAMAKOSHI A., TANAKA T.,
RA   TERPSTRA P., TOGNONI A., TOSATO V., UCHIYAMA S., VANDENBOL M.,
RA   VANNIER F., VASSAROTTI A., VIARI A., WAMBUTT R., WEDLER E., WEDLER H.,
RA   WEITZENEGGER T., WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K.,
RA   YASUMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E.,
RA   YOSHIKAWA H., DANCHIN A.;
RL   NATURE 390:249-256(1997).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
RL   SUBMITTED (NOV-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z99115; E1183584; -.
SQ   SEQUENCE   875 AA;  98885 MW;  C0AD7428 CRC32;

  Query Match           59.3%;  Score 48;  DB 9;  Length 875;
  Best Local Similarity  45.5%;  Pred. No. 2.18e+00;
  Matches     5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    231 INSVNPTGQSY 241
          ::  :|||:||
Qy      2 VKEADPTGHSY 12


RESULT   14
ID   O30764         PRELIMINARY;   PRT;  4340 AA.
AC   O30764;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   POLYKETIDE SYNTHASE MODULES 1 AND 2.
GN   NIDA1.
OS   STREPTOMYCES CAELESTIS.
OC   EUBACTERIA; FIRMICUTES; ACTINOMYCETES; STREPTOMYCETES;
OC   STREPTOMYCETACEAE; STREPTOMYCES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRRL-2821;
RA   KAKAVAS S.J., KATZ L., STASSI D.;
RL   J. BACTERIOL. 179:7515-7522(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRRL-2821;
RA   KAKAVAS S., STASSI D.;
RL   SUBMITTED (JUL-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF016585; G2558838; -.
DR   PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
KW   TRANSFERASE.
SQ   SEQUENCE   4340 AA;  457589 MW;  659B2ECB CRC32;

  Query Match           59.3%;  Score 48;  DB 9;  Length 4340;
  Best Local Similarity  60.0%;  Pred. No. 2.18e+00;
  Matches     6;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db   3618 VRAADPAGHP 3627
          |: |||:||:
Qy      2 VKEADPTGHS 11


RESULT   15
ID   Q09717         PRELIMINARY;   PRT;   215 AA.
AC   Q09717;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   RAN/SPI1 BINDING PROTEIN.
GN   SBP1.
OS   SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   HAYASHI N., SAZER S., NISHIMOTO T.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; D86381; G1408521; -.
SQ   SEQUENCE   215 AA;  24150 MW;  1D19BCB7 CRC32;

  Query Match           58.0%;  Score 47;  DB 1;  Length 215;
  Best Local Similarity  41.7%;  Pred. No. 3.84e+00;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db    173 DVSEGEPTAETF 184
          || |:::|: ::
Qy      1 DVKEADPTGHSY 12


RESULT   16
ID   O15479         PRELIMINARY;   PRT;   313 AA.
AC   O15479;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   MAGE-B2.
GN   MAGE-B2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   LURQUIN C.;
RL   SUBMITTED (OCT-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U93163; G2459680; -.
SQ   SEQUENCE   313 AA;  34739 MW;  9543930D CRC32;

  Query Match           58.0%;  Score 47;  DB 2;  Length 313;
  Best Local Similarity  33.3%;  Pred. No. 3.84e+00;
  Matches     4;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db    167 ELNKVNPNGHTY 178
          :::  :| ||:|
Qy      1 DVKEADPTGHSY 12


RESULT   17
ID   P74837         PRELIMINARY;   PRT;   108 AA.
AC   P74837;
DT   01-FEB-1997 (TREMBLREL. 02, CREATED)
DT   01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT   01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
DE   NAPHTHALENESULFONIC ACID DIOXYGENASE FERREDOXIN SUBUNIT.
OS   SPHINGOMONAS SP.
OC   EUBACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; ZYMOMONAS GROUP;
OC   SPHINGOMONAS.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BN6;
RA   CONRADT D., KLEIN J., MATTES R.;
RL   SUBMITTED (AUG-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U65001; G1513105; -.
KW   DIOXYGENASE.
SQ   SEQUENCE   108 AA;  11595 MW;  5B7478FD CRC32;

  Query Match           56.8%;  Score 46;  DB 9;  Length 108;
  Best Local Similarity  55.6%;  Pred. No. 6.72e+00;
  Matches     5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db     12 DVKDGEPVG 20
          |||:::| |
Qy      1 DVKEADPTG 9


RESULT   18
ID   P95779         PRELIMINARY;   PRT;   198 AA.
AC   P95779;
DT   01-MAY-1997 (TREMBLREL. 03, CREATED)
DT   01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT   01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE   DTDP-4-KETO-L-RHAMNOSE REDUCTASE.
GN   RMLC.
OS   STREPTOCOCCUS MUTANS.
OC   PROKARYOTA; FIRMICUTES; COCCI; STREPTOCOCCACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=XC;
```

```
RA   TSUKIOKA Y., YAMASHITA Y., NAKANO Y., OHO T., KOGA T.;
RL   SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=XC;
RA   TSUKIOKA Y., YAMASHITA Y., NAKANO Y., OHO T., KOGA T.;
RL   SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; D78182; G1813346; -.
SQ   SEQUENCE   198 AA;  22364 MW;  8C148392 CRC32;

  Query Match           56.8%;  Score 46;  DB 9;  Length 198;
  Best Local Similarity  50.0%;  Pred. No. 6.72e+00;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     99 DLREGDSFGHVY 110
          |::|:|: || |
Qy      1 DVKEADPTGHSY 12


RESULT   19
ID   O17311         PRELIMINARY;   PRT;   272 AA.
AC   O17311;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   20S PROTEASOME BETA2 SUBUNIT (EC 3.4.99.46).
GN   BETA2_DM.
OS   DROSOPHILA MELANOGASTER (FRUIT FLY).
OC   EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   SMYTH K.A., BELOTE J.M.;
RL   SUBMITTED (NOV-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF025791; G2582504; -.
DR   PROSITE; PS00854; PROTEASOME_B; 1.
KW   PROTEASOME; HYDROLASE; PROTEASE.
SQ   SEQUENCE   272 AA;  29895 MW;  3E6F291D CRC32;

  Query Match           56.8%;  Score 46;  DB 3;  Length 272;
  Best Local Similarity  75.0%;  Pred. No. 6.72e+00;
  Matches     6;  Conservative    2;  Mismatches   0;  Indels   0;  Gaps   0;

Db    194 VRDADPTG 201
          |::|||||
Qy      2 VKEADPTG 9


RESULT   20
ID   O17312         PRELIMINARY;   PRT;   272 AA.
AC   O17312;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   20S PROTEASOME BETA2 SUBUNIT (EC 3.4.99.46).
GN   BETA2_DM.
OS   DROSOPHILA MELANOGASTER (FRUIT FLY).
OC   EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   SMYTH K.A., BELOTE J.M.;
RL   SUBMITTED (SEP-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF025792; G2582506; -.
DR   PROSITE; PS00854; PROTEASOME_B; 1.
KW   PROTEASOME; HYDROLASE; PROTEASE.
SQ   SEQUENCE   272 AA;  29883 MW;  5688BBDA CRC32;

  Query Match           56.8%;  Score 46;  DB 3;  Length 272;
  Best Local Similarity  75.0%;  Pred. No. 6.72e+00;
  Matches     6;  Conservative    2;  Mismatches   0;  Indels   0;  Gaps   0;

Db    194 VRDADPTG 201
          |::|||||
Qy      2 VKEADPTG 9


RESULT   21
ID   Q41681         PRELIMINARY;   PRT;   317 AA.
AC   Q41681;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   1-AMINOCYCLOPROPANE-1-CARBOXYLATE
DE   OXIDASE HOMOLOG.
OS   PHASEOLUS AUREUS (MUNG BEAN) (VIGNA RADIATA).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
OC   FABACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HYPOCOTYL;
RX   MEDLINE; 94339795.
RA   KIM W.T., YANG S.F.;
RL   PLANTA 194:223-229(1994).
DR   EMBL; U06046; G458338; -.
SQ   SEQUENCE   317 AA;  35773 MW;  F02E6272 CRC32;

  Query Match           56.8%;  Score 46;  DB 8;  Length 317;
  Best Local Similarity  54.5%;  Pred. No. 6.72e+00;
  Matches     6;  Conservative    3;  Mismatches   2;  Indels   0;  Gaps   0;

Db    266 VKESDETSQVY 276
          |||:| |:: |
Qy      2 VKEADPTGHSY 12


RESULT   22
ID   P92966         PRELIMINARY;   PRT;   356 AA.
AC   P92966;
DT   01-MAY-1997 (TREMBLREL. 03, CREATED)
DT   01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT   01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE   SPLICING FACTOR.
GN   RSP41.
OS   ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC   CAPPARALES; CRUCIFERAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   LOPATO S., WAIGMANN E., BARTA A.;
RL   SUBMITTED (DEC-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; X99436; E258268; -.
SQ   SEQUENCE   356 AA;  41297 MW;  7F090666 CRC32;

  Query Match           56.8%;  Score 46;  DB 8;  Length 356;
  Best Local Similarity  45.5%;  Pred. No. 6.72e+00;
  Matches     5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    166 VKDDDSRGNGY 176
          ||: |: |::|
Qy      2 VKEADPTGHSY 12


RESULT   23
ID   Q51532         PRELIMINARY;   PRT;   382 AA.
AC   Q51532;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   PILU.
GN   PILU.
OS   PSEUDOMONAS AERUGINOSA.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   PSEUDOMONADACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE; 91285432.
```

```
RA   WHITCHURCH C.B., HOBBS M., LIVINGSTON S.P., KRISHNAPILLAI V.,
RA   MATTICK J.S.;
RL   GENE 101:33-44(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RA   WHITCHURCH C.B., MATTICK J.S.;
RL   SUBMITTED (MAR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; L27667; G443686; -.
SQ   SEQUENCE   382 AA;  42533 MW;  6F4A16C2 CRC32;

  Query Match           56.8%;  Score 46;  DB 9;  Length 382;
  Best Local Similarity  25.0%;  Pred. No. 6.72e+00;
  Matches     3;  Conservative    6;  Mismatches   3;  Indels   0;  Gaps   0;

Db    370 EITDDDPAGRRF 381
          :: : ||:|: :
Qy      1 DVKEADPTGHSY 12


RESULT   24
ID   Q40565         PRELIMINARY;   PRT;   439 AA.
AC   Q40565;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   RIBULOSE BISPHOSPHATE CARBOXYLASE ACTIVASE.
GN   RCA.
OS   NICOTIANA TABACUM (COMMON TOBACCO).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC   SOLANALES; SOLANACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SR1; TISSUE=LEAF;
RA   RODERMEL S., QIAN J.;
RL   SUBMITTED (AUG-1992) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z14980; G19990; -.
SQ   SEQUENCE   439 AA;  48343 MW;  0E004E3C CRC32;

  Query Match           56.8%;  Score 46;  DB 8;  Length 439;
  Best Local Similarity  41.7%;  Pred. No. 6.72e+00;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db     60 EEKDADPKKQTY 71
          : |:|||  ::|
Qy      1 DVKEADPTGHSY 12


RESULT   25
ID   Q24044         PRELIMINARY;   PRT;   526 AA.
AC   Q24044;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   FIZZY (FZY).
GN   FZY.
OS   DROSOPHILA MELANOGASTER (FRUIT FLY).
OC   EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95247821.
RA   DAWSON I.A., ROTH S., ARTAVANIS-TSAKONAS S.;
RL   J. CELL BIOL. 129:725-737(1995).
DR   EMBL; U22419; G1109772; -.
DR   FLYBASE; FBgn0001086; fzy.
SQ   SEQUENCE   526 AA;  57080 MW;  122FC6EB CRC32;

  Query Match           56.8%;  Score 46;  DB 3;  Length 526;
  Best Local Similarity  70.0%;  Pred. No. 6.72e+00;
  Matches     7;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db    461 VKQADLTGHT 470
          ||:|| |||:

Qy      2 VKEADPTGHS 11


RESULT   26
ID   Q05664         PRELIMINARY;   PRT;   543 AA.
AC   Q05664;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   TRANSCRIPTION FACTOR.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ALPHA S288;
RA   MANNHAUPT G., VETTER I., SCHWARZLOSE C., MITZEL S., FELDMANN H.;
RL   YEAST 0:0-0(0).
DR   EMBL; X91067; G984187; -.
SQ   SEQUENCE   543 AA;  60305 MW;  25845EE2 CRC32;

  Query Match           56.8%;  Score 46;  DB 1;  Length 543;
  Best Local Similarity  45.5%;  Pred. No. 6.72e+00;
  Matches     5;  Conservative    1;  Mismatches   5;  Indels   0;  Gaps   0;

Db    285 VKDNGPINHVY 295
          ||:  |  | |
Qy      2 VKEADPTGHSY 12


RESULT   27
ID   Q08224         PRELIMINARY;   PRT;   551 AA.
AC   Q08224;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   CHROMOSOME XV READING FRAME ORF YOL055C.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   FELDMANN H., MANNHAUPT G., VETTER I.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   ANSORGE W., BENES V., RECHMANN S., SCHWAGER C., TEODORU C., VOSS H.,
RA   WIEMANN S.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   MIPS;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z74797; E251864; -.
SQ   SEQUENCE   551 AA;  61269 MW;  DDAD22AB CRC32;

  Query Match           56.8%;  Score 46;  DB 1;  Length 551;
  Best Local Similarity  45.5%;  Pred. No. 6.72e+00;
  Matches     5;  Conservative    1;  Mismatches   5;  Indels   0;  Gaps   0;

Db    285 VKDNGPINHVY 295
          ||:  |  | |
Qy      2 VKEADPTGHSY 12


RESULT   28
ID   Q08975         PRELIMINARY;   PRT;   551 AA.
AC   Q08975;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   ORF YPL258C.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RA   MESSENGUY F., DUBOIS E., VIERENDEELS F., SCHERENS B.;
RL   SUBMITTED (JUN-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MIPS;
RL   SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z73614; E246962; -.
SQ   SEQUENCE   551 AA;  61334 MW;  E3E8364D CRC32;

  Query Match           56.8%;  Score 46;  DB 1;  Length 551;
  Best Local Similarity  45.5%;  Pred. No. 6.72e+00;
  Matches    5;  Conservative    1;  Mismatches   5;  Indels   0;  Gaps   0;

Db    285 VKDNGPINHVY 295
          ||:  |  | |
Qy      2 VKEADPTGHSY 12


RESULT   29
ID   Q63615         PRELIMINARY;   PRT;   597 AA.
AC   Q63615;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE   VACUOLAR PROTEIN SORTING HOMOLOG R-VPS33A.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   PEVSNER J., HSU S.C., HYDE P.S., SCHELLER R.H.;
RL   GENE 183:7-14(1996).
DR   EMBL; U35244; G1477468; -.
SQ   SEQUENCE   597 AA;  67513 MW;  7469FB50 CRC32;

  Query Match           56.8%;  Score 46;  DB 10;  Length 597;
  Best Local Similarity  58.3%;  Pred. No. 6.72e+00;
  Matches    7;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db    477 DVNEQNPTDISY 488
          ||:| :||  ||
Qy      1 DVKEADPTGHSY 12


RESULT   30
ID   Q38231         PRELIMINARY;   PRT;   628 AA.
AC   Q38231;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   UNIDENTIFIED ORF34.
OS   BACTERIOPHAGE BIL67.
OC   VIRIDAE; DS-DNA NONENVELOPED VIRUSES; SIPHOVIRIDAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95111629.
RA   SCHOULER C., EHRLICH S.D., CHOPIN M.C.;
RL   MICROBIOLOGY 140:3061-3069(1994).
DR   EMBL; L33769; G522256; -.
SQ   SEQUENCE   628 AA;  72844 MW;  7292CDBD CRC32;

  Query Match           56.8%;  Score 46;  DB 7;  Length 628;
  Best Local Similarity  36.4%;  Pred. No. 6.72e+00;
  Matches    4;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    596 IKDENPQNQSY 606
          :|: :|  :||
Qy      2 VKEADPTGHSY 12


RESULT   31
ID   Q93539         PRELIMINARY;   PRT;   702 AA.
AC   Q93539;
DT   01-FEB-1997 (TREMBLREL. 02, CREATED)
DT   01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT   01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
DE   F20D1.7.
OS   CAENORHABDITIS ELEGANS.
OC   EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   BURTON J.;
RL   SUBMITTED (NOV-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,
RA   COULSON A., CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A.,
RA   FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
RA   JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
RA   LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B.,
RA   O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A.,
RA   SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A., SONNHAMMER E.,
RA   STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M.,
RA   VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA   WILKINSON-SPROAT J., WOHLDMAN P.;
RL   NATURE 368:32-38(1994).
DR   EMBL; Z78542; E259712; -.
SQ   SEQUENCE   702 AA;  79995 MW;  8818ED6B CRC32;

  Query Match           56.8%;  Score 46;  DB 3;  Length 702;
  Best Local Similarity  36.4%;  Pred. No. 6.72e+00;
  Matches    4;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    413 IRQTATVGHSY 423
          ::::  ||||
Qy      2 VKEADPTGHSY 12


RESULT   32
ID   Q12866         PRELIMINARY;   PRT;   999 AA.
AC   Q12866;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   CELLULAR PROTO-ONCOGENE (C-MER) PRECURSOR (C-MER).
GN   C-MER.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PERIPHERAL BLOOD LEUKOCYTES;
RX   MEDLINE; 94368701.
RA   GRAHAM D.K., DAWSON T.L., MULLANEY D.L., SNODGRASS H.R., EARP H.S.;
RL   CELL GROWTH DIFFER. 5:647-657(1994).
RN   [2]
RP   ERRATUM.
RA   GRAHAM D.K., DAWSON T.L., MULLANEY D.L., SNODGRASS H.R., EARP H.S.;
RL   CELL GROWTH DIFFER. 5:1022-1022(1994).
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. THE
CC       EXTRACELLULAR DOMAINS CONTAINS 2 IG-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; U08023; G505665; -.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   PROTO-ONCOGENE; TRANSMEMBRANE; REPEAT; IMMUNOGLOBULIN FOLD; SIGNAL.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    999       CELLULAR PROTO-ONCOGNE (C-MER).
FT   DOMAIN       19    501       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    502    518       POTENTIAL.
FT   DOMAIN      519    999       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       90    182       IG-LIKE DOMAIN.
FT   DOMAIN      198    266       IG-LIKE DOMAIN.
```

```
FT   DOMAIN      283    372       FIBRONECTIN TYPE-III.
FT   DOMAIN      376    480       FIBRONECTIN TYPE-III.
SQ   SEQUENCE   999 AA;  110391 MW;  20274ED6 CRC32;

  Query Match           56.8%;  Score 46;  DB 2;  Length 999;
  Best Local Similarity  63.6%;  Pred. No. 6.72e+00;
  Matches     7;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    320 QVKEADPLGNG 330
          :|||||| |::
Qy      1 DVKEADPTGHS 11


RESULT   33
ID   Q59278         PRELIMINARY;   PRT;  1187 AA.
AC   Q59278;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE   ENDOXYLANASE (EC 3.2.1.8) (ENDO-1,4-BETA-XYLANASE)
DE   (1,4-BETA-D-XYLAN XYLANOHYDROLASE).
GN   XYNC.
OS   CELLULOMONAS FIMI.
OC   PROKARYOTA; FIRMICUTES; IRREGULAR ASPOROGENOUS RODS; CORYNEFORM GROUP.
RN   [1]
RP   SEQUENCE OF 1-352 FROM N.A.
RA   CLARKE J.H., DAVIDSON K., GILBERT H.J., HAZLEWOOD G.P.;
RL   SUBMITTED (AUG-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   CLARKE J.H.;
RL   SUBMITTED (AUG-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
CC       LINKAGES IN XYLANS.
DR   EMBL; Z50866; E212269; -.
KW   XYLAN DEGRADATION; HYDROLASE; GLYCOSIDASE.
SQ   SEQUENCE   1187 AA;  125378 MW;  92B3994A CRC32;

  Query Match           56.8%;  Score 46;  DB 9;  Length 1187;
  Best Local Similarity  75.0%;  Pred. No. 6.72e+00;
  Matches     6;  Conservative    2;  Mismatches   0;  Indels   0;  Gaps   0;

Db   1046 TDPTGRSY 1053
          :||||:||
Qy      5 ADPTGHSY 12


RESULT   34
ID   O34568         PRELIMINARY;   PRT;    46 AA.
AC   O34568;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   YBBU PROTEIN.
GN   YBBU.
OS   BACILLUS SUBTILIS.
OC   PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
RA   AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
RA   BORRISS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
RA   BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
RA   CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
RA   DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,
RA   ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,
RA   FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,
RA   GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,
RA   GUISEPPI G., GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A.,
RA   HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,
RA   JORIS B., KARAMATA D., KASAHARA Y., KLAERR-BLANCHARD M., KLEIN C.,
RA   KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,
RA   KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
RA   LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,
RA   MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,
RA   NOONE D., O'REILLY M., OGAWA K., OGIWARA A., OUDEGA B., PARK S.H.,
RA   PARRO V., POHL T.M., PORTETELLE D., PORWOLLIK S., PRESCOTT A.M.,
RA   PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,
RA   RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAIE Y.,
RA   SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,
RA   SEKIGUCHI J., SEKOWSKA A., SEROR S.J., SERROR P., SHIN B.S., SOLDO B.,
RA   SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,
RA   TAKEUCHI M., TAMAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,
RA   TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,
RA   VIARI A., WAMBUTT R., WEDLER E., WEDLER H., WEITZENEGGER T.,
RA   WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,
RA   YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;
RL   NATURE 390:249-256(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
RL   SUBMITTED (NOV-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; Z99104; E1182112; -.
DR   EMBL; Z99105; E1182130; -.
SQ   SEQUENCE   46 AA;  5370 MW;  60D493CA CRC32;

  Query Match           55.6%;  Score 45;  DB 9;  Length 46;
  Best Local Similarity  54.5%;  Pred. No. 1.16e+01;
  Matches     6;  Conservative    1;  Mismatches   4;  Indels   0;  Gaps   0;

Db    14 IKEVDGTGPDY 24
          :|| | ||  |
Qy     2 VKEADPTGHSY 12


RESULT   35
ID   Q61740         PRELIMINARY;   PRT;    86 AA.
AC   Q61740;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   INTEGRIN ALPHA 4 (CD49D) (INTEGRIN ALPHA-4 SUBUNIT) (FRAGMENT).
GN   ITGA4 OR VLA-4.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95290094.
RA   DE MEIRSMAN C., SCHOLLEN E., JASPERS M., ONGENA K., MATTHIJS G.,
RA   MARYNEN P., CASSIMAN J.J.;
RL   DNA CELL BIOL. 13:743-754(1994).
DR   EMBL; L20788; G309417; -.
DR   MGD; MGI:96603; ITGA4.
KW   INTEGRIN.
FT   NON_TER      86     86
SQ   SEQUENCE   86 AA;  9255 MW;  C6D01AE6 CRC32;

  Query Match           55.6%;  Score 45;  DB 10;  Length 86;
  Best Local Similarity 100.0%;  Pred. No. 1.16e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    29 PTGHSY 34
          ||||||
Qy     7 PTGHSY 12


Search completed: Tue Apr  7 08:42:16 1998
Job time : 12 secs.
```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:41:29 1998; MasPar time 2.18 Seconds
138.083 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-4
Description: (1-12) from 5541104.pep
Perfect Score: 81
Sequence: 1 DVKEADPTGHSY 12

Scoring table: PAM 150
Gap 15

Searched: 69112 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 23.518; Variance 22.821; scale 1.031

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	81	100.0	309	1	MAG1_HUMAN	MELANOMA-ASSOCIATED AN	4.36e-10
2	74	91.4	234	1	MAG8_HUMAN	MELANOMA-ASSOCIATED AN	7.50e-08
3	74	91.4	315	1	MAG9_HUMAN	MELANOMA-ASSOCIATED AN	7.50e-08
4	72	88.9	369	1	MAGA_HUMAN	MELANOMA-ASSOCIATED AN	3.15e-07
5	71	87.7	319	1	MAGB_HUMAN	MELANOMA-ASSOCIATED AN	6.43e-07
6	62	76.5	317	1	MAG4_HUMAN	MELANOMA-ASSOCIATED AN	3.14e-04
7	57	70.4	347	1	MAGX_HUMAN	MELANOMA-ASSOCIATED AN	8.07e-03
8	53	65.4	314	1	MAG6_HUMAN	MELANOMA-ASSOCIATED AN	9.61e-02
9	52	64.2	417	1	F16P_ARATH	FRUCTOSE-1,6-BISPHOSPH	1.75e-01
10	50	61.7	314	1	MAG3_HUMAN	MELANOMA-ASSOCIATED AN	5.69e-01
11	50	61.7	407	1	F16P_PEA	FRUCTOSE-1,6-BISPHOSPH	5.69e-01
12	50	61.7	835	1	FASD_ECOLI	OUTER MEMBRANE USHER P	5.69e-01
13	49	60.5	411	1	F16P_BRANA	FRUCTOSE-1,6-BISPHOSPH	1.01e+00
14	49	60.5	417	1	CBPC_HUMAN	MAST CELL CARBOXYPEPTI	1.01e+00
15	49	60.5	3898	1	POLG_BVDVS	GENOME POLYPROTEIN.	1.01e+00
16	48	59.3	268	1	RAD_RAT	GTP-BINDING PROTEIN RA	1.79e+00
17	48	59.3	269	1	RAD_HUMAN	GTP-BINDING PROTEIN RA	1.79e+00
18	48	59.3	866	1	YCBS_ECOLI	HYPOTHETICAL OUTER MEM	1.79e+00
19	48	59.3	873	1	PC1_HUMAN	PLASMA-CELL MEMBRANE G	1.79e+00
20	47	58.0	197	1	KGUA_PIG	GUANYLATE KINASE (EC 2	3.12e+00
21	47	58.0	443	1	GLNA_BACCE	GLUTAMINE SYNTHETASE (	3.12e+00
22	47	58.0	443	1	GLNA_BACSU	GLUTAMINE SYNTHETASE (	3.12e+00
23	47	58.0	445	1	GLNA_LACDE	GLUTAMINE SYNTHETASE (	3.12e+00
24	47	58.0	446	1	GLNA_STAAU	GLUTAMINE SYNTHETASE (	3.12e+00
25	47	58.0	497	1	PEN3_ADE12	PENTON PROTEIN (VIRION	3.12e+00
26	47	58.0	633	1	XRCC_HUMAN	DNA-REPAIR PROTEIN XRC	3.12e+00
27	46	56.8	147	1	PA24_BUNMU	PHOSPHOLIPASE A2, BETA	5.40e+00
28	46	56.8	583	1	ARA2_ECOLI	ARSENICAL PUMP-DRIVING	5.40e+00
29	46	56.8	583	1	ARA1_ECOLI	ARSENICAL PUMP-DRIVING	5.40e+00
30	45	55.6	503	1	VP57_BDV	57 KD PROTEIN (P57).	9.26e+00
31	45	55.6	581	1	PBP2_NEIGO	PENICILLIN-BINDING PRO	9.26e+00
32	45	55.6	581	1	PBP2_NEIME	PENICILLIN-BINDING PRO	9.26e+00
33	45	55.6	725	1	MYT1_HUMAN	MYELIN TRANSCRIPTION F	9.26e+00
34	45	55.6	1033	1	T1R1_ECOLI	TYPE I RESTRICTION ENZ	9.26e+00
35	44	54.3	22	1	YHV4_LACHE	HYPOTHETICAL PROTEIN I	1.57e+01
36	44	54.3	98	1	Y037_MYCTU	VERY HYPOTHETICAL 11.3	1.57e+01
37	44	54.3	222	1	THIE_BACSU	THIAMIN-PHOSPHATE PYRO	1.57e+01
38	44	54.3	261	1	143A_VICFA	14-3-3-LIKE PROTEIN A	1.57e+01
39	44	54.3	377	1	NUEM_HUMAN	NADH-UBIQUINONE OXIDOR	1.57e+01
40	44	54.3	469	1	GLN1_RHIME	GLUTAMINE SYNTHETASE I	1.57e+01
41	44	54.3	469	1	GLN1_RHILV	GLUTAMINE SYNTHETASE I	1.57e+01
42	44	54.3	488	1	SUOX_RAT	SULFITE OXIDASE PRECUR	1.57e+01
43	44	54.3	488	1	SUOX_HUMAN	SULFITE OXIDASE PRECUR	1.57e+01
44	44	54.3	539	1	CCMM_SYNP7	CARBON DIOXIDE CONCENT	1.57e+01
45	44	54.3	629	1	HDF2_YEAST	HIGH AFFINITY DNA-BIND	1.57e+01
46	44	54.3	789	1	ACOX_YEAST	PUTATIVE ACONITASE IN	1.57e+01
47	44	54.3	826	1	VIL1_HUMAN	VILLIN.	1.57e+01
48	44	54.3	826	1	VIL1_MOUSE	VILLIN.	1.57e+01
49	44	54.3	1173	1	TSP1_XENLA	THROMBOSPONDIN 1 PRECU	1.57e+01
50	43	53.1	176	1	RL5_HALMA	50S RIBOSOMAL PROTEIN	2.63e+01
51	43	53.1	228	1	RL1_THETH	50S RIBOSOMAL PROTEIN	2.63e+01
52	43	53.1	290	1	YG1D_YEAST	HYPOTHETICAL 31.7 KD P	2.63e+01
53	43	53.1	293	1	YS31_MYCTU	HYPOTHETICAL 33.0 KD P	2.63e+01
54	43	53.1	301	1	SC14_KLULA	SEC14 CYTOSOLIC FACTOR	2.63e+01
55	43	53.1	326	1	YKX2_CAEEL	HYPOTHETICAL 34.6 KD P	2.63e+01
56	43	53.1	362	1	YCHF_HAEIN	PROBABLE GTP-BINDING P	2.63e+01
57	43	53.1	416	1	CBPB_CANFA	CARBOXYPEPTIDASE B PRE	2.63e+01
58	43	53.1	506	1	Z157_HUMAN	ZINC FINGER PROTEIN 15	2.63e+01
59	43	53.1	559	1	PPBI_MOUSE	ALKALINE PHOSPHATASE,	2.63e+01
60	43	53.1	669	1	COGU_HUMAN	MATRIX METALLOPROTEINA	2.63e+01
61	43	53.1	684	1	CNG1_MOUSE	CGMP-GATED CATION CHAN	2.63e+01
62	43	53.1	686	1	CNG1_HUMAN	CGMP-GATED CATION CHAN	2.63e+01
63	43	53.1	690	1	CNG1_BOVIN	CGMP-GATED CATION CHAN	2.63e+01
64	43	53.1	691	1	CNG1_CANFA	CGMP-GATED CATION CHAN	2.63e+01
65	43	53.1	796	1	ABAA_EMENI	REGULATORY PROTEIN ABA	2.63e+01
66	43	53.1	878	1	YB9X_YEAST	HYPOTHETICAL TRP-ASP R	2.63e+01
67	43	53.1	2504	1	FAS_HUMAN	FATTY ACID SYNTHASE (E	2.63e+01
68	43	53.1	3224	1	N358_HUMAN	NUCLEAR PORE COMPLEX P	2.63e+01
69	43	53.1	3396	1	POLG_DEN1S	GENOME POLYPROTEIN (CO	2.63e+01
70	42	51.9	145	1	ANG3_MOUSE	ANGIOGENIN-3 PRECURSOR	4.37e+01
71	42	51.9	150	1	RPB8_HUMAN	DNA-DIRECTED RNA POLYM	4.37e+01
72	42	51.9	156	1	YLXS_BACSU	HYPOTHETICAL 17.6 KD P	4.37e+01
73	42	51.9	156	1	VPG_PLRVW	PUTATIVE GENOME-LINKED	4.37e+01
74	42	51.9	156	1	VPG_PLRV	PUTATIVE GENOME-LINKED	4.37e+01
75	42	51.9	156	1	VPG_PLRV1	PUTATIVE GENOME-LINKED	4.37e+01
76	42	51.9	156	1	VPG_PLRVR	PUTATIVE GENOME-LINKED	4.37e+01
77	42	51.9	197	1	KGUA_MOUSE	GUANYLATE KINASE (EC 2	4.37e+01
78	42	51.9	283	1	FOLD_BACSU	METHYLENETETRAHYDROFOL	4.37e+01
79	42	51.9	292	1	TF_RABIT	TISSUE FACTOR PRECURSO	4.37e+01
80	42	51.9	342	1	YN9A_YEAST	HYPOTHETICAL 37.9 KD P	4.37e+01
81	42	51.9	347	1	PYRC_SALTY	DIHYDROOROTASE (EC 3.5	4.37e+01
82	42	51.9	363	1	YCHF_ECOLI	PROBABLE GTP-BINDING P	4.37e+01
83	42	51.9	379	1	ILEU_HORSE	LEUKOCYTE ELASTASE INH	4.37e+01
84	42	51.9	392	1	CGL_CAEEL	PUTATIVE CYSTATHIONINE	4.37e+01
85	42	51.9	417	1	PVR_HUMAN	POLIOVIRUS RECEPTOR PR	4.37e+01
86	42	51.9	516	1	ACHD_BOVIN	ACETYLCHOLINE RECEPTOR	4.37e+01
87	42	51.9	555	1	WETA_EMENI	REGULATORY PROTEIN WET	4.37e+01
88	42	51.9	621	1	VP40_HSVBC	CAPSID PROTEIN P40 (CO	4.37e+01
89	42	51.9	699	1	EFG_AGRTU	ELONGATION FACTOR G (E	4.37e+01
90	42	51.9	700	1	EFG_AQUPY	ELONGATION FACTOR G (E	4.37e+01
91	42	51.9	747	1	GUND_CELFI	ENDOGLUCANASE D PRECUR	4.37e+01
92	42	51.9	773	1	MAK5_YEAST	ATP-DEPENDENT RNA HELI	4.37e+01
93	42	51.9	799	1	ZFX1_MOUSE	ZINC FINGER X-CHROMOSO	4.37e+01
94	42	51.9	839	1	ZFX2_MOUSE	ZINC FINGER X-CHROMOSO	4.37e+01
95	42	51.9	991	1	BMP1_MOUSE	BONE MORPHOGENETIC PRO	4.37e+01
96	42	51.9	1073	1	RESA_PLAFF	RING-INFECTED ERYTHROC	4.37e+01

97	42	51.9	1132	1	YKK5_YEAST	HYPOTHETICAL 125.6 KD	4.37e+01
98	42	51.9	1139	1	HMW1_MYCGE	CYTADHERENCE HIGH MOLE	4.37e+01
99	41	50.6	366	1	CD44_BOVIN	CD44 ANTIGEN PRECURSOR	7.17e+01
100	41	50.6	407	1	CG1E_CHICK	G1/S-SPECIFIC CYCLIN E	7.17e+01

ALIGNMENTS

```
RESULT    1
ID   MAG1_HUMAN     STANDARD;      PRT;   309 AA.
AC   P43355; O00346;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 1 (MAGE-1 ANTIGEN) (ANTIGEN MZ2-E).
GN   MAGEA1 OR MAGE1 OR MAGE1A.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92086861.
RA   VAN DER BRUGGEN P., TRAVERSARI C., CHOMEZ P., LURQUIN C., DE PLAEN E.,
RA   VAN DEN EYNDE B., KNUTH A., BOON T.;
RL   SCIENCE 254:1643-1647(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKIN;
RX   MEDLINE; 94311935.
RA   DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   GLOECKNER G., RUMP A., NORDSIEK G., HINZMANN B., KIOSCHIS P.,
RA   HEISS N., POUSTKA A., BAUER D., DRESCHER B., KNOB A., ROSENTHAL A.;
RL   SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [4]
RP   MUTAGENESIS.
RC   TISSUE=BLOOD;
RX   MEDLINE; 94157413.
RA   GAUGLER B., VAN DEN EYNDE B., VAN DER BRUGGEN P., ROMERO P.,
RA   GAFORIO J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;
RL   J. EXP. MED. 179:921-930(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC       CYTOLYTIC T LYMPHOCYTES.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES. NEVER EXPRESSED IN KIDNEY TUMORS, LEUKEMIAS AND
CC       LYMPHOMAS.
CC   -!- POLYMORPHISM: THE VARIANT AT POSITION 32 LIKELY REPRESENTS A
CC       POLYMORPHISM OF THE MAG-1 GENE.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; M77481; G416115; -.
DR   EMBL; U82672; G2078527; -.
DR   MIM; 300016; -.
KW   ANTIGEN; MULTIGENE FAMILY; POLYMORPHISM; TUMOR ANTIGEN.
FT   VARIANT      32     32       T -> A.
FT   DOMAIN       33     36       POLY-SER.
FT   MUTAGEN     163    163       D->A: ABOLISHES HLA-A1 BINDING.
FT   MUTAGEN     169    169       Y->A: ABOLISHES HLA-A1 BINDING.
FT   CONFLICT     72     72       R -> Q (IN REF. 3).
SQ   SEQUENCE   309 AA;  34342 MW;  E6CB1300 CRC32;

  Query Match          100.0%;  Score 81;  DB 1;  Length 309;
  Best Local Similarity 100.0%;  Pred. No. 4.36e-10;
  Matches    12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    158 DVKEADPTGHSY 169
          ||||||||||||
Qy      1 DVKEADPTGHSY 12
```

```
RESULT    2
ID   MAG8_HUMAN     STANDARD;      PRT;   234 AA.
AC   P43361;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 8 (MAGE-8 ANTIGEN).
GN   MAGEA8 OR MAGE8.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; U10693; G533526; -.
KW   ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT   DOMAIN       40     43       POLY-SER.
SQ   SEQUENCE   234 AA;  25197 MW;  D4931BC3 CRC32;

  Query Match          91.4%;  Score 74;  DB 1;  Length 234;
  Best Local Similarity 83.3%;  Pred. No. 7.50e-08;
  Matches    10;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    168 DVKEVDPAGHSY 179
          |||| ||:||||
Qy      1 DVKEADPTGHSY 12
```

```
RESULT    3
ID   MAG9_HUMAN     STANDARD;      PRT;   315 AA.
AC   P43362; Q92910;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 9 (MAGE-9 ANTIGEN).
GN   MAGEA9 OR MAGE9.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   TIMMS K.M., BONDESON M.L., ANSARI-LARI M.A., LAGERSTEDT K.,
RA   NELSON D.L., PETTERSSON U., GIBBS R.A.;
RL   SUBMITTED (SEP-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
```

```
DR   EMBL; U10694; G533528; -.
DR   EMBL; U66083; G1519285; -.
KW   ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT   DOMAIN       34     37       POLY-GLU.
FT   DOMAIN       87     90       POLY-GLU.
SQ   SEQUENCE   315 AA;  35088 MW;  7DC3228E CRC32;

  Query Match           91.4%;  Score 74;  DB 1;  Length 315;
  Best Local Similarity  83.3%;  Pred. No. 7.50e-08;
  Matches    10;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    164 DVKEVDPAGHSY 175
          |||| ||:||||
Qy      1 DVKEADPTGHSY 12


RESULT    4
ID   MAGA_HUMAN     STANDARD;      PRT;   369 AA.
AC   P43363;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 10 (MAGE-10 ANTIGEN).
GN   MAGEA10 OR MAGE10.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
DR   EMBL; U10685; G533511; -.
KW   ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT   DOMAIN       54     62       POLY-SER.
SQ   SEQUENCE   369 AA;  40766 MW;  D11E1870 CRC32;

  Query Match           88.9%;  Score 72;  DB 1;  Length 369;
  Best Local Similarity  83.3%;  Pred. No. 3.15e-07;
  Matches    10;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    190 DVKEVDPTGHSF 201
          |||| ||||||:
Qy      1 DVKEADPTGHSY 12


RESULT    5
ID   MAGB_HUMAN     STANDARD;      PRT;   319 AA.
AC   P43364;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 11 (MAGE-11 ANTIGEN).
GN   MAGEA11 OR MAGE11.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; U10686; G533513; -.
KW   ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
SQ   SEQUENCE   319 AA;  35536 MW;  E3DBEDEF CRC32;

  Query Match           87.7%;  Score 71;  DB 1;  Length 319;
  Best Local Similarity  83.3%;  Pred. No. 6.43e-07;
  Matches    10;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    168 DVKEVDPTSHSY 179
          |||| |||:|||
Qy      1 DVKEADPTGHSY 12


RESULT    6
ID   MAG4_HUMAN     STANDARD;      PRT;   317 AA.
AC   P43358;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 4 (MAGE-4 ANTIGEN) (MAGE-X2) (MAGE-41).
GN   MAGEA4 OR MAGE4.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKIN;
RX   MEDLINE; 94311935.
RA   DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95369706.
RA   IMAI Y., SHICHIJO S., YAMADA A., KATAYAMA T., YANO H., ITOH K.;
RL   GENE 160:287-290(1995).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY WITH
CC       MAGE-1.
DR   EMBL; U10687; G533515; -.
DR   EMBL; U10688; G533517; -.
DR   EMBL; U10340; G499124; -.
DR   EMBL; D32077; G1125018; -.
KW   ANTIGEN; MULTIGENE FAMILY; POLYMORPHISM; TUMOR ANTIGEN.
FT   DOMAIN       41     44       POLY-SER.
FT   VARIANT     173    173       T -> A.
FT   CONFLICT    307    307       E -> Q (IN REF. 2).
SQ   SEQUENCE   317 AA;  34929 MW;  3CE38AF9 CRC32;

  Query Match           76.5%;  Score 62;  DB 1;  Length 317;
  Best Local Similarity  66.7%;  Pred. No. 3.14e-04;
  Matches     8;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;
```

```
Db    166 DVKEVDPTSNTY 177
          |||| |||:::|
Qy      1 DVKEADPTGHSY 12


RESULT    7
ID   MAGX_HUMAN     STANDARD;      PRT;   347 AA.
AC   P43366;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN XP (MAGE-XP ANTIGEN).
GN   MAGEB1 OR MAGEL1 OR MAGEXP.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE; 95281581.
RA   MUSCATELLI F., WALKER A.P., DE PLAEN E., STAFFORD A.N., MONACO A.P.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 92:4987-4991(1995).
CC   -!- TISSUE SPECIFICITY: EXPRESSED ONLY IN TESTIS.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; X82539; G608993; -.
DR   MIM; 600619; -.
KW   ANTIGEN; MULTIGENE FAMILY.
SQ   SEQUENCE   347 AA;  39152 MW;  A041BAB2 CRC32;

  Query Match           70.4%;  Score 57;  DB 1;  Length 347;
  Best Local Similarity  50.0%;  Pred. No. 8.07e-03;
  Matches     6;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db    164 DLKEDNPSSHTY 175
          |:|| :|:::|:|
Qy      1 DVKEADPTGHSY 12


RESULT    8
ID   MAG6_HUMAN     STANDARD;      PRT;   314 AA.
AC   P43360;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 6 (MAGE-6 ANTIGEN) (MAGE3B).
GN   MAGEA6 OR MAGE6.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   DE PLAEN E., ARDEN K., TRAVERSARI C., GAFORIO J.J., SZIKORA J.-P.,
RA   DE SMET C., BRASSEUR F., VAN DER BRUGGEN P., LETHE B., LURQUIN C.,
RA   BRASSEUR R., CHOMEZ P., DE BACKER O., CAVENEE W., BOON T.;
RL   IMMUNOGENETICS 40:360-369(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKIN;
RX   MEDLINE; 94311935.
RA   DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95369706.
RA   IMAI Y., SHICHIJO S., YAMADA A., KATAYAMA T., YANO H., ITOH K.;
RL   GENE 160:287-290(1995).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN TUMOR
CC       OR ASPECTS OF TUMOR PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY TO
CC       MAGE-3.
DR   EMBL; U10691; G533523; -.
DR   EMBL; U10339; G499122; -.
DR   EMBL; D32076; G1125016; -.
KW   ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT   DOMAIN       40     43       POLY-SER.
SQ   SEQUENCE   314 AA;  34891 MW;  B7125E97 CRC32;

  Query Match           65.4%;  Score 53;  DB 1;  Length 314;
  Best Local Similarity  50.0%;  Pred. No. 9.61e-02;
  Matches     6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db    165 ELMEVDPIGHVY 176
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT    9
ID   F16P_ARATH     STANDARD;      PRT;   417 AA.
AC   P25851;
DT   01-MAY-1992 (REL. 22, CREATED)
DT   01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   FRUCTOSE-1,6-BISPHOSPHATASE, CHLOROPLAST PRECURSOR (EC 3.1.3.11)
DE   (D-FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE).
GN   FBP.
OS   ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC   CAPPARALES; CRUCIFERAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91329733.
RA   HORSNELL P.R., RAINES C.A.;
RL   PLANT MOL. BIOL. 17:185-186(1991).
CC   -!- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE + H(2)O =
CC       D-FRUCTOSE 6-PHOSPHATE + ORTHOPHOSPHATE.
CC   -!- PATHWAY: THE CHLOROPLAST ISOZYME TAKES PART IN THE REGENERATION OF
CC       RIBULOSE BISPHOSPHATE IN THE PHOTOSYNTHETIC CARBON REDUCTION
CC       CYCLE (CALVIN CYCLE).
CC   -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- INDUCTION: LIGHT ACTIVATION THROUGH PH CHANGES, MG(2+) LEVELS
CC       AND ALSO BY LIGHT-MODULATED REDUCTION OF ESSENTIAL DISULPHIDE
CC       GROUPS VIA THE FERREDOXIN-THIOREDOXIN F SYSTEM (BY SIMILARITY).
CC   -!- IN PLANTS THERE ARE TWO FBPASE ISOZYMES: ONE IN THE CYTOSOL AND
CC       THE OTHER IN THE CHLOROPLAST.
DR   EMBL; X58148; G11242; -.
DR   PIR; S16582; S16582.
DR   HSSP; P00636; 1FBH.
DR   PROSITE; PS00124; FBPASE; 1.
KW   HYDROLASE; CARBOHYDRATE METABOLISM; MULTIGENE FAMILY; CHLOROPLAST;
KW   TRANSIT PEPTIDE; CALVIN CYCLE.
FT   TRANSIT       1     59       CHLOROPLAST (POTENTIAL).
FT   CHAIN        60    417       FRUCTOSE-1,6-BISPHOSPHATASE.
FT   ACT_SITE    359    359       BY SIMILARITY.
FT   DISULFID    233    238       REDOX-ACTIVE (LIGHT-MODULATED) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   417 AA;  45178 MW;  9A30B20C CRC32;

  Query Match           64.2%;  Score 52;  DB 1;  Length 417;
  Best Local Similarity  50.0%;  Pred. No. 1.75e-01;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    314 DLKDPGPTGKPY 325
          |:|:: ||| :|
Qy      1 DVKEADPTGHSY 12


RESULT   10
ID   MAG3_HUMAN     STANDARD;      PRT;   314 AA.
AC   P43357;
```

```
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MELANOMA-ASSOCIATED ANTIGEN 3 (MAGE-3 ANTIGEN) (ANTIGEN MZ2-D).
GN   MAGEA3 OR MAGE3.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS.
RC   TISSUE=BLOOD;
RX   MEDLINE; 94157413.
RA   GAUGLER B., VAN DEN EYNDE B., VAN DER BRUGGEN P., ROMERO P.,
RA   GAFORIO J.J., DE PLAEN E., LETHE B., BRASSEUR F., BOON T.;
RL   J. EXP. MED. 179:921-930(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKIN;
RX   MEDLINE; 94311935.
RA   DING M., BECK R.J., KELLER C.J., FENTON R.G.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 202:549-555(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC       CYTOLYTIC T LYMPHOCYTES.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA. NEVER EXPRESSED IN KIDNEY TUMORS,
CC       LEUKEMIAS AND LYMPHOMAS.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; U03735; G468826; -.
KW   ANTIGEN; MULTIGENE FAMILY; TUMOR ANTIGEN.
FT   DOMAIN       40     43       POLY-SER.
FT   MUTAGEN     170    170       D->A: ABOLISHES HLA-A1 BINDING.
FT   MUTAGEN     176    176       Y->A: ABOLISHES HLA-A1 BINDING.
SQ   SEQUENCE   314 AA;  34747 MW;  AC557A64 CRC32;

  Query Match           61.7%;  Score 50;  DB 1;  Length 314;
  Best Local Similarity  50.0%;  Pred. No. 5.69e-01;
  Matches     6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db    165 ELMEVDPIGHLY 176
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT   11
ID   F16P_PEA       STANDARD;      PRT;   407 AA.
AC   P46275; Q37263;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   FRUCTOSE-1,6-BISPHOSPHATASE, CHLOROPLAST PRECURSOR (EC 3.1.3.11)
DE   (D-FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE).
GN   FBP.
OS   PISUM SATIVUM (GARDEN PEA).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
OC   FABACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. GIANT; TISSUE=LEAF;
RA   HAHN T.R., DONG S.M., RHIM J.H.;
RL   SUBMITTED (JAN-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE OF 27-407 FROM N.A.
RC   STRAIN=CV. LINCOLN; TISSUE=LEAF;
RX   MEDLINE; 94297517.
RA   CARRASCO J.L., CHUECA A., PRADO F.E., HERMOSO R., LAZARO J.J.,
RA   RAMOS J.L., SAHRAWY M., LOPEZ GORGE J.;
RL   PLANTA 193:494-501(1994).
CC   -!- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE + H(2)O =
CC       D-FRUCTOSE 6-PHOSPHATE + ORTHOPHOSPHATE.
CC   -!- PATHWAY: THE CHLOROPLAST ISOZYME TAKES PART IN THE REGENERATION OF
CC       RIBULOSE BISPHOSPHATE IN THE PHOTOSYNTHETIC CARBON REDUCTION
CC       CYCLE (CALVIN CYCLE).
CC   -!- SUBUNIT: HOMOTETRAMER.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST STROMA.
CC   -!- INDUCTION: LIGHT ACTIVATION THROUGH PH CHANGES, MG(2+) LEVELS
CC       AND ALSO BY LIGHT-MODULATED REDUCTION OF ESSENTIAL DISULFIDE
CC       GROUPS VIA THE FERREDOXIN-THIOREDOXIN F SYSTEM (BY SIMILARITY).
CC   -!- IN PLANTS THERE ARE TWO FBPASE ISOZYMES: ONE IN THE CYTOSOL AND
CC       THE OTHER IN THE CHLOROPLAST.
DR   EMBL; L34806; G609561; -.
DR   EMBL; X68826; G20717; -.
DR   PROSITE; PS00124; FBPASE; 1.
KW   HYDROLASE; CARBOHYDRATE METABOLISM; MULTIGENE FAMILY; CHLOROPLAST;
KW   TRANSIT PEPTIDE; CALVIN CYCLE.
FT   TRANSIT       1     50       CHLOROPLAST (POTENTIAL).
FT   CHAIN        51    407       FRUCTOSE-1,6-BISPHOSPHATASE.
FT   ACT_SITE    349    349       BY SIMILARITY.
FT   DISULFID    223    228       REDOX-ACTIVE (LIGHT-MODULATED) (BY
FT                                SIMILARITY).
FT   CONFLICT     82     82       G -> P (IN REF. 2).
FT   CONFLICT    160    160       A -> P (IN REF. 2).
FT   CONFLICT    247    247       A -> I (IN REF. 2).
FT   CONFLICT    282    282       E -> K (IN REF. 2).
SQ   SEQUENCE   407 AA;  44511 MW;  DD67CDEC CRC32;

  Query Match           61.7%;  Score 50;  DB 1;  Length 407;
  Best Local Similarity  50.0%;  Pred. No. 5.69e-01;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    304 DLKEPGPSGKPY 315
          |:||: |:| :|
Qy      1 DVKEADPTGHSY 12


RESULT   12
ID   FASD_ECOLI     STANDARD;      PRT;   835 AA.
AC   P46000;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   OUTER MEMBRANE USHER PROTEIN FASD PRECURSOR.
GN   FASD.
OS   ESCHERICHIA COLI.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=987;
RX   MEDLINE; 94148769.
RA   SCHIFFERLI D.M., ALRUTZ M.A.;
RL   J. BACTERIOL. 176:1099-1110(1994).
CC   -!- FUNCTION: INVOLVED IN THE EXPORT AND ASSEMBLY OF THE 987P
CC       FIMBRIAE SUBUNITS ACROSS THE OUTER MEMBRANE.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.
CC   -!- SIMILARITY: TO OTHER FIMBRIAL EXPORT USHER PROTEINS.
DR   EMBL; L22659; G437336; -.
DR   EMBL; U50547; G1381551; -.
DR   PROSITE; PS01151; FIMBRIAL_USHER; 1.
KW   OUTER MEMBRANE; TRANSMEMBRANE; FIMBRIA; TRANSPORT; SIGNAL.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    835       OUTER MEMBRANE USHER PROTEIN FASD.
FT   DISULFID    810    834       POTENTIAL.
SQ   SEQUENCE   835 AA;  92354 MW;  D8FBDD31 CRC32;

  Query Match           61.7%;  Score 50;  DB 1;  Length 835;
  Best Local Similarity  50.0%;  Pred. No. 5.69e-01;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    324 NIKEADGSEHSF 335
          ::|||| : ||:
Qy      1 DVKEADPTGHSY 12
```

```
RESULT   13
ID   F16P_BRANA     STANDARD;      PRT;   411 AA.
AC   Q07204;
DT   01-OCT-1994 (REL. 30, CREATED)
DT   01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   FRUCTOSE-1,6-BISPHOSPHATASE, CHLOROPLAST PRECURSOR (EC 3.1.3.11)
DE   (D-FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE).
GN   FBP.
OS   BRASSICA NAPUS (RAPE).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC   CAPPARALES; CRUCIFERAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94120014.
RA   RODRIGUEZ-SUAREZ R.J., WOLOSIUK R.A.;
RL   PLANT PHYSIOL. 103:1453-1454(1993).
CC   -!- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE + H(2)O =
CC       D-FRUCTOSE 6-PHOSPHATE + ORTHOPHOSPHATE.
CC   -!- PATHWAY: THE CHLOROPLAST ISOZYME TAKES PART IN THE REGENERATION OF
CC       RIBULOSE BISPHOSPHATE IN THE PHOTOSYNTHETIC CARBON REDUCTION
CC       CYCLE (CALVIN CYCLE).
CC   -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST STROMA.
CC   -!- INDUCTION: LIGHT ACTIVATION THROUGH PH CHANGES, MG(2+) LEVELS
CC       AND ALSO BY LIGHT-MODULATED REDUCTION OF ESSENTIAL DISULFIDE
CC       GROUPS VIA THE FERREDOXIN-THIOREDOXIN F SYSTEM (BY SIMILARITY).
CC   -!- IN PLANTS THERE ARE TWO FBPASE ISOZYMES: ONE IN THE CYTOSOL AND
CC       THE OTHER IN THE CHLOROPLAST.
DR   EMBL; L15303; G289367; -.
DR   HSSP; P00636; 1FRP.
DR   PROSITE; PS00124; FBPASE; 1.
KW   HYDROLASE; CARBOHYDRATE METABOLISM; MULTIGENE FAMILY; CHLOROPLAST;
KW   TRANSIT PEPTIDE; CALVIN CYCLE.
FT   TRANSIT       1     53       CHLOROPLAST (BY SIMILARITY).
FT   CHAIN        54    411       FRUCTOSE-1,6-BISPHOSPHATASE.
FT   ACT_SITE    353    353       BY SIMILARITY.
FT   DISULFID    227    232       REDOX-ACTIVE (LIGHT-MODULATED) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   411 AA;  44446 MW;  BFBD2BCD CRC32;

  Query Match           60.5%;  Score 49;  DB 1;  Length 411;
  Best Local Similarity  41.7%;  Pred. No. 1.01e+00;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db    308 DLKDPGPSGKPY 319
          |:|:: |:| :|
Qy      1 DVKEADPTGHSY 12


RESULT   14
ID   CBPC_HUMAN     STANDARD;      PRT;   417 AA.
AC   P15088;
DT   01-APR-1990 (REL. 14, CREATED)
DT   01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   MAST CELL CARBOXYPEPTIDASE A PRECURSOR (EC 3.4.17.1) (MC-CPA)
DE   (CARBOXYPEPTIDASE A3).
GN   CPA3.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RX   MEDLINE; 90083291.
RA   REYNOLDS D.S., GURLEY D.S., STEVENS R.L., SUGARBAKER D.J.,
RA   AUSTEN K.F., SERAFIN W.E.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 86:9480-9484(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MAST CELLS;
RX   MEDLINE; 92105393.
RA   REYNOLDS D.S., GURLEY D.S., AUSTEN K.F.;
RL   J. CLIN. INVEST. 89:273-282(1992).
RN   [3]
RP   SEQUENCE OF 110-417 FROM N.A.
RX   MEDLINE; 92333165.
RA   NATSUAKI M., STEWART C.B., VANDERSLICE P., SCHWARTZ L.B.,
RA   NATSUAKI M., WINTROUB B.U., RUTTER W.J., GOLDSTEIN S.M.;
RL   J. INVEST. DERMATOL. 99:138-145(1992).
RN   [4]
RP   SEQUENCE OF 110-137.
RX   MEDLINE; 89214692.
RA   GOLDSTEIN S.M., KAEMPFER C.E., KEALEY J.T., WINTROUB B.U.;
RL   J. CLIN. INVEST. 83:1630-1636(1989).
CC   -!- CATALYTIC ACTIVITY: PEPTIDYL-L-AMINO ACID + H(2)O = PEPTIDE +
CC       L-AMINO ACID.
CC   -!- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M14; ALSO KNOWN AS THE
CC       ZINC CARBOXYPEPTIDASE FAMILY.
DR   EMBL; M27717; G179934; -.
DR   EMBL; M73720; G187442; -.
DR   EMBL; M73716; G187442; JOINED.
DR   EMBL; M73717; G187442; JOINED.
DR   EMBL; M73718; G187442; JOINED.
DR   EMBL; M73719; G187442; JOINED.
DR   EMBL; S40234; E65321; ALT_SEQ.
DR   PIR; A43929; A43929.
DR   HSSP; P09955; 1PBA.
DR   MIM; 114851; -.
DR   PROSITE; PS00132; CARBOXYPEPT_ZN_1; 1.
DR   PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
KW   HYDROLASE; CARBOXYPEPTIDASE; ZINC; ZYMOGEN; SIGNAL.
FT   SIGNAL        1     15
FT   PROPEP       16    109       ACTIVATION PEPTIDE.
FT   CHAIN       110    417       MAST CELL CARBOXYPEPTIDASE A.
FT   METAL       176    176       ZINC (BY SIMILARITY).
FT   METAL       179    179       ZINC (BY SIMILARITY).
FT   METAL       304    304       ZINC (BY SIMILARITY).
FT   ACT_SITE    378    378       NUCLEOPHILE (BY SIMILARITY).
FT   DISULFID    173    186       BY SIMILARITY.
FT   DISULFID    245    268       BY SIMILARITY.
SQ   SEQUENCE   417 AA;  48700 MW;  848CEC99 CRC32;

  Query Match           60.5%;  Score 49;  DB 1;  Length 417;
  Best Local Similarity  66.7%;  Pred. No. 1.01e+00;
  Matches     8;  Conservative    0;  Mismatches   4;  Indels   0;  Gaps   0;

Db    105 DVKEDIPGRHSY 116
          |||| |  |||
Qy      1 DVKEADPTGHSY 12


RESULT   15
ID   POLG_BVDVS     STANDARD;      PRT;  3898 AA.
AC   Q01499;
DT   01-JUL-1993 (REL. 26, CREATED)
DT   01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   GENOME POLYPROTEIN.
OS   BOVINE VIRAL DIARRHEA VIRUS (STRAIN SD-1) (BVDV) (MUCOSAL DISEASE
OS   VIRUS).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; FLAVIVIRIDAE;
OC   PESTIVIRUSES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93079889.
RA   DENG R., BROCK K.V.;
RL   VIROLOGY 191:867-869(1992).
CC   -!- FUNCTION: PESTIVIRUS P80 (P125) MAY BE A BIFUNCTIONAL PROTEIN
CC       WITH HELICASE AND PROTEASE ACTIVITY.
CC   -!- PTM: GP116 GIVES RISE TO GP62 AND GP53; GP62 IN TURN YIELDS GP48
CC       AND GP25.
CC   -!- SIMILARITY: TO THE HOG CHOLERA VIRUS GENOME POLYPROTEIN.
```

```
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S31.
DR   EMBL; M96751; G289508; -.
DR   PIR; A44217; A44217.
DR   PROSITE; PS00531; RNASE_T2_2; UNKNOWN_1.
KW   POLYPROTEIN; GLYCOPROTEIN; HELICASE; SERINE PROTEASE; HYDROLASE.
FT   CHAIN         1   ?270       P20 (30KD).
FT   CHAIN      ?271  ?1063       GP116/GP62-GP53 (GLYCOPROTEIN).
FT   CHAIN         ?      ?       GP125/GP54-GP80.
FT   CHAIN         ?   3898       GP133/GP58-GP75.
FT   DOMAIN      690    755       CYS-RICH.
FT   ACT_SITE   1658   1658       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   1695   1695       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   1752   1752       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   CARBOHYD    272    272       POTENTIAL.
FT   CARBOHYD    281    281       POTENTIAL.
FT   CARBOHYD    296    296       POTENTIAL.
FT   CARBOHYD    335    335       POTENTIAL.
FT   CARBOHYD    365    365       POTENTIAL.
FT   CARBOHYD    370    370       POTENTIAL.
FT   CARBOHYD    413    413       POTENTIAL.
FT   CARBOHYD    487    487       POTENTIAL.
FT   CARBOHYD    597    597       POTENTIAL.
FT   CARBOHYD    809    809       POTENTIAL.
FT   CARBOHYD    878    878       POTENTIAL.
FT   CARBOHYD    922    922       POTENTIAL.
FT   CARBOHYD    990    990       POTENTIAL.
FT   CARBOHYD   1357   1357       POTENTIAL.
FT   CARBOHYD   1419   1419       POTENTIAL.
FT   CARBOHYD   1451   1451       POTENTIAL.
FT   CARBOHYD   1713   1713       POTENTIAL.
FT   CARBOHYD   2134   2134       POTENTIAL.
FT   CARBOHYD   2217   2217       POTENTIAL.
FT   CARBOHYD   2494   2494       POTENTIAL.
FT   CARBOHYD   2682   2682       POTENTIAL.
FT   CARBOHYD   2751   2751       POTENTIAL.
FT   CARBOHYD   2891   2891       POTENTIAL.
FT   CARBOHYD   2988   2988       POTENTIAL.
FT   CARBOHYD   3688   3688       POTENTIAL.
FT   CARBOHYD   3777   3777       POTENTIAL.
FT   CARBOHYD   3793   3793       POTENTIAL.
SQ   SEQUENCE   3898 AA;  437800 MW;  A562145C CRC32;

  Query Match           60.5%;  Score 49;  DB 1;  Length 3898;
  Best Local Similarity  58.3%;  Pred. No. 1.01e+00;
  Matches     7;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db   1132 DVVKADPGGQGY 1143
          ||  ||| |::|
Qy      1 DVKEADPTGHSY 12


RESULT   16
ID   RAD_RAT        STANDARD;      PRT;   268 AA.
AC   P55043;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GTP-BINDING PROTEIN RAD (RAS ASSOCIATED WITH DIABETES) (RAD1).
GN   RRAD OR RAD.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   RISHI A.K., GULAMHUSSEIN A., STEELE M.P.;
RL   SUBMITTED (DEC-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- SIMILARITY: BELONGS TO THE RAD/GEM FAMILY OF GTP-BINDING
CC       PROTEINS.
DR   EMBL; U12187; G595473; -.
KW   GTP-BINDING.
FT   NP_BIND      59     66       GTP (BY SIMILARITY).
FT   NP_BIND     108    112       GTP (BY SIMILARITY).
FT   NP_BIND     163    166       GTP (BY SIMILARITY).
SQ   SEQUENCE   268 AA;  29053 MW;  CE4C8DC7 CRC32;

  Query Match           59.3%;  Score 48;  DB 1;  Length 268;
  Best Local Similarity  55.6%;  Pred. No. 1.79e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     80 EAEAAGHTY 88
          ||:::||:|
Qy      4 EADPTGHSY 12


RESULT   17
ID   RAD_HUMAN      STANDARD;      PRT;   269 AA.
AC   P55042;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   GTP-BINDING PROTEIN RAD (RAS ASSOCIATED WITH DIABETES) (RAD1).
GN   RRAD OR RAD.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKELETAL MUSCLE;
RX   MEDLINE; 94069319.
RA   REYNET C., KAHN C.R.;
RL   SCIENCE 262:1441-1444(1993).
CC   -!- TISSUE SPECIFICITY: SKELETAL AND CARDIAC MUSCLE, LUNG, LESSER
CC       AMOUNTS IN PLACENTA AND KIDNEY. DETECTED IN ADIPOSE TISSUE.
CC       OVEREXPRESSED IN MUSCLE OF TYPE II DIABETIC HUMANS.
CC   -!- SIMILARITY: BELONGS TO THE RAD/GEM FAMILY OF GTP-BINDING
CC       PROTEINS.
DR   EMBL; L24564; G439603; -.
DR   MIM; 179503; -.
KW   GTP-BINDING.
FT   NP_BIND      59     66       GTP (BY SIMILARITY).
FT   NP_BIND     108    112       GTP (BY SIMILARITY).
FT   NP_BIND     164    167       GTP (BY SIMILARITY).
SQ   SEQUENCE   269 AA;  29262 MW;  201C9665 CRC32;

  Query Match           59.3%;  Score 48;  DB 1;  Length 269;
  Best Local Similarity  55.6%;  Pred. No. 1.79e+00;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     80 EAEAAGHTY 88
          ||:::||:|
Qy      4 EADPTGHSY 12


RESULT   18
ID   YCBS_ECOLI     STANDARD;      PRT;   866 AA.
AC   P75857;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL OUTER MEMBRANE USHER PROTEIN IN PEPN-PYRD INTERGENIC
DE   REGION PRECURSOR.
GN   YCBS.
OS   ESCHERICHIA COLI.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RA   BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL   SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: INVOLVED IN THE EXPORT AND ASSEMBLY OF A FIMBRIAL
CC       SUBUNIT ACROSS THE OUTER MEMBRANE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER FIMBRIAL EXPORT USHER PROTEINS.
```

```
DR   EMBL; AE000196; G1787172; -.
DR   ECOGENE; EG13711; YCBS.
DR   PROSITE; PS01151; FIMBRIAL_USHER; 1.
KW   HYPOTHETICAL PROTEIN; OUTER MEMBRANE; TRANSMEMBRANE; FIMBRIA;
KW   TRANSPORT; SIGNAL.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    866       HYPOTHETICAL OUTER MEMBRANE USHER PROTEIN
FT                                YCBS.
SQ   SEQUENCE   866 AA;  95241 MW;  62A077F2 CRC32;

  Query Match           59.3%;  Score 48;  DB 1;  Length 866;
  Best Local Similarity  50.0%;  Pred. No. 1.79e+00;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    334 EIKEADGSVNSY 345
          ::|||| : :||
Qy      1 DVKEADPTGHSY 12


RESULT   19
ID   PC1_HUMAN      STANDARD;      PRT;   873 AA.
AC   P22413;
DT   01-AUG-1991 (REL. 19, CREATED)
DT   01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 (ALKALINE PHOSPHODIESTERASE I
DE   (EC 3.1.4.1) / NUCLEOTIDE PYROPHOSPHATASE (EC 3.6.1.9) (NPPASE)).
GN   PC1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91009202.
RA   BUCKLEY M.F., LOVELAND K.A., MCKINSTRY W.J., GARSON O.M., GODING J.W.;
RL   J. BIOL. CHEM. 265:17506-17511(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92246539.
RA   FUNAKOSHI I., KATO H., HORIE K., YANO T., HORI Y., KOBAYASHI H.,
RA   INOUE T., SUZUKI H., FUKUI S., TSUKAHARA M., KAJII T.,
RA   YAMASHINA I.;
RL   ARCH. BIOCHEM. BIOPHYS. 295:180-187(1992).
CC   -!- FUNCTION: MAY HAVE A ROLE IN THE REGULATION OF N-GLYCOSYLATION.
CC   -!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES
CC       SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED
CC       OLIGO-NUCLEOTIDES.
CC   -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.
CC   -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN PLASMA CELLS AND ALSO IN A NUMBER
CC       OF NON-LYMPHOID TISSUES, INCLUDING THE DISTAL CONVOLUTED TUBULE
CC       OF THE KIDNEY, CHONDROCYTES, AND EPIDIDYMIS.
CC   -!- SIMILARITY: CONTAINS TWO TANDEM COPIES OF A SOMATOMEDIN-B TYPE
CC       DOMAIN.
DR   EMBL; M57736; G189650; -.
DR   EMBL; D12485; G219945; -.
DR   EMBL; D12485; G219944; ALT_INIT.
DR   PIR; A39216; A39216.
DR   MIM; 173335; -.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   GLYCOPROTEIN; TRANSMEMBRANE; DUPLICATION; SIGNAL-ANCHOR; HYDROLASE.
FT   DOMAIN        1     24       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     25     45       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN       46    873       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       52     92       SOMATOMEDIN-B LIKE.
FT   DOMAIN       93    136       SOMATOMEDIN-B LIKE.
FT   CARBOHYD    127    127       POTENTIAL.
FT   CARBOHYD    233    233       POTENTIAL.
FT   CARBOHYD    289    289       POTENTIAL.
FT   CARBOHYD    425    425       POTENTIAL.
FT   CARBOHYD    533    533       POTENTIAL.
FT   CARBOHYD    591    591       POTENTIAL.
FT   CARBOHYD    648    648       POTENTIAL.
FT   CARBOHYD    679    679       POTENTIAL.
FT   CARBOHYD    696    696       POTENTIAL.
SQ   SEQUENCE   873 AA;  99929 MW;  580583CD CRC32;

  Query Match           59.3%;  Score 48;  DB 1;  Length 873;
  Best Local Similarity  66.7%;  Pred. No. 1.79e+00;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db    322 EPDSSGHSY 330
          |:|::||||
Qy      4 EADPTGHSY 12


RESULT   20
ID   KGUA_PIG       STANDARD;      PRT;   197 AA.
AC   P31006;
DT   01-JUL-1993 (REL. 26, CREATED)
DT   01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   GUANYLATE KINASE (EC 2.7.4.8) (GMP KINASE).
GN   GUK1 OR GUK.
OS   SUS SCROFA (PIG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93238695.
RA   ZSCHOCKE P.D., SCHILTZ E., SCHULZ G.E.;
RL   EUR. J. BIOCHEM. 213:263-269(1993).
CC   -!- FUNCTION: ESSENTIAL FOR RECYCLING GMP AND INDIRECTLY, CGMP.
CC   -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SIMILARITY: TO OTHER GUANYLATE KINASES.
DR   PIR; S23776; KIPGGU.
DR   HSSP; P15454; 1GKY.
DR   PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR   PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
KW   TRANSFERASE; KINASE; ATP-BINDING; ACETYLATION.
FT   INIT_MET      0      0
FT   MOD_RES       1      1       ACETYLATION.
FT   NP_BIND      10     17       ATP (BY SIMILARITY).
SQ   SEQUENCE   197 AA;  21789 MW;  7F9FAF05 CRC32;

  Query Match           58.0%;  Score 47;  DB 1;  Length 197;
  Best Local Similarity  54.5%;  Pred. No. 3.12e+00;
  Matches     6;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    187 EIKKAQATGHS 197
          ::| |::||||
Qy      1 DVKEADPTGHS 11


RESULT   21
ID   GLNA_BACCE     STANDARD;      PRT;   443 AA.
AC   P19064;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   GLUTAMINE SYNTHETASE (EC 6.3.1.2) (GLUTAMATE--AMMONIA LIGASE).
GN   GLNA.
OS   BACILLUS CEREUS.
OC   PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.
RC   STRAIN=IFO 3131;
RX   MEDLINE; 90036764.
RA   NAKANO Y., KATO C., TANAKA E., KIMURA K., HORIKOSHI K.;
RL   J. BIOCHEM. 106:209-215(1989).
CC   -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +
CC       ORTHOPHOSPHATE.
CC   -!- ENZYME REGULATION: NOT REGULATED BY POST-TRANSLATIONAL
```

```
CC       MODIFICATION AND NOT SUBJECT TO FEEDBACK INHIBITION.
CC   -!- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC       HEXAGON.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
DR   EMBL; D00513; G216273; -.
DR   PIR; JU0075; AJBSQU.
DR   HSSP; P06201; 1LGR.
DR   PROSITE; PS00180; GLNA_1; 1.
DR   PROSITE; PS00181; GLNA_ATP; 1.
DR   PROSITE; PS00182; GLNA_ADENYLATION; 1.
KW   NITROGEN FIXATION; LIGASE.
FT   INIT_MET      0      0
SQ   SEQUENCE   443 AA;  50064 MW;  5617AB71 CRC32;

  Query Match           58.0%;  Score 47;  DB 1;  Length 443;
  Best Local Similarity  33.3%;  Pred. No. 3.12e+00;
  Matches     4;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    332 EVRSVDPAANPY 343
          :|:  ||::::|
Qy      1 DVKEADPTGHSY 12


RESULT   22
ID   GLNA_BACSU     STANDARD;      PRT;   443 AA.
AC   P12425;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   GLUTAMINE SYNTHETASE (EC 6.3.1.2) (GLUTAMATE--AMMONIA LIGASE).
GN   GLNA.
OS   BACILLUS SUBTILIS.
OC   PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89138001.
RA   STRAUCH M.A., ARONSON A.I., BROWN S.W., SCHREIER H.J.,
RA   SONENSHEIN A.L.;
RL   GENE 71:257-265(1988).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC   STRAIN=PCI219;
RX   MEDLINE; 89232679.
RA   NAKANO Y., TANAKA E., KATO C., KIMURA K., HORIKOSHI K.;
RL   FEMS MICROBIOL. LETT. 48:81-86(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   BORCHERT S., KLEIN C., PIKSA B., HAMMELMANN M., ENTIAN K.D.;
RL   SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +
CC       ORTHOPHOSPHATE.
CC   -!- ENZYME REGULATION: NOT REGULATED BY POST-TRANSLATIONAL
CC       MODIFICATION AND NOT SUBJECT TO FEEDBACK INHIBITION.
CC   -!- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC       HEXAGON.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
DR   EMBL; M22811; G142986; -.
DR   EMBL; D00854; G216396; -.
DR   EMBL; U66480; G1750111; -.
DR   PIR; JT0392; AJBSQS.
DR   PIR; A48312; A48312.
DR   HSSP; P06201; 1LGR.
DR   SUBTILIST; BG10425; GLNA.
DR   PROSITE; PS00180; GLNA_1; 1.
DR   PROSITE; PS00181; GLNA_ATP; 1.
DR   PROSITE; PS00182; GLNA_ADENYLATION; 1.
KW   NITROGEN FIXATION; LIGASE.
FT   INIT_MET      0      0
FT   VARIANT       9      9       E -> V (IN STRAIN PCI219).
FT   VARIANT      42     42       G -> E (IN STRAIN PCI219).
FT   VARIANT     252    252       N -> D (IN STRAIN PCI219).
FT   VARIANT     258    258       F -> Y (IN STRAIN PCI219).
SQ   SEQUENCE   443 AA;  50147 MW;  D0D84747 CRC32;

  Query Match           58.0%;  Score 47;  DB 1;  Length 443;
  Best Local Similarity  33.3%;  Pred. No. 3.12e+00;
  Matches     4;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    332 EVRSVDPAANPY 343
          :|:  ||::::|
Qy      1 DVKEADPTGHSY 12


RESULT   23
ID   GLNA_LACDE     STANDARD;      PRT;   445 AA.
AC   P45627;
DT   01-NOV-1995 (REL. 32, CREATED)
DT   01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   GLUTAMINE SYNTHETASE (EC 6.3.1.2) (GLUTAMATE--AMMONIA LIGASE).
GN   GLNA.
OS   LACTOBACILLUS DELBRUECKII (SUBSP. BULGARICUS).
OC   PROKARYOTA; FIRMICUTES; REGULAR ASPOROGENOUS ROD; LACTOBACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93073924.
RA   ISHINO Y., MORGENTHALER P., HOTTINGER H., SOELL D.;
RL   APPL. ENVIRON. MICROBIOL. 58:3165-3169(1992).
CC   -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +
CC       ORTHOPHOSPHATE.
CC   -!- ENZYME REGULATION: DOES NOT SEEM TO BE REGULATED BY ADENYLATION.
CC   -!- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC       HEXAGON (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
DR   EMBL; D10020; G216749; -.
DR   PROSITE; PS00180; GLNA_1; 1.
DR   PROSITE; PS00181; GLNA_ATP; 1.
KW   NITROGEN FIXATION; LIGASE.
SQ   SEQUENCE   445 AA;  50133 MW;  F923BB07 CRC32;

  Query Match           58.0%;  Score 47;  DB 1;  Length 445;
  Best Local Similarity  33.3%;  Pred. No. 3.12e+00;
  Matches     4;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db    335 EMRSTDPTANPY 346
          ::: :|||:::|
Qy      1 DVKEADPTGHSY 12


RESULT   24
ID   GLNA_STAAU     STANDARD;      PRT;   446 AA.
AC   Q59812;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   GLUTAMINE SYNTHETASE (EC 6.3.1.2) (GLUTAMATE--AMMONIA LIGASE) (GS).
GN   GLNA.
OS   STAPHYLOCOCCUS AUREUS.
OC   PROKARYOTA; FIRMICUTES; COCCI; MICROCOCCACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BB270 / AS63;
RA   STRANDEN A.M., BERGER-BAECHI B.;
RL   SUBMITTED (DEC-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +
CC       ORTHOPHOSPHATE.
CC   -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) =
CC       ADP + GLUTAMINE + ORTHOPHOSPHATE.
CC   -!- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC       HEXAGON (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
DR   EMBL; X76490; E214721; -.
DR   PROSITE; PS00180; GLNA_1; 1.
DR   PROSITE; PS00181; GLNA_ATP; 1.
KW   NITROGEN FIXATION; LIGASE.
SQ   SEQUENCE   446 AA;  50840 MW;  830BFD81 CRC32;
```

```
  Query Match           58.0%;  Score 47;  DB 1;  Length 446;
  Best Local Similarity  33.3%;  Pred. No. 3.12e+00;
  Matches    4;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    335 EVRSVDPAANPY 346
          :|:  ||:::::|
Qy      1 DVKEADPTGHSY 12


RESULT   25
ID   PEN3_ADE12     STANDARD;      PRT;   497 AA.
AC   P36716;
DT   01-JUN-1994 (REL. 29, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   PENTON PROTEIN (VIRION COMPONENT III) (PENTON BASE PROTEIN).
GN   PIII.
OS   HUMAN ADENOVIRUS TYPE 12.
OC   VIRIDAE; DS-DNA NONENVELOPED VIRUSES; ADENOVIRIDAE; MASTADENOVIRUSES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94076430.
RA   SPRENGEL J., SCHMITZ B., HEUSS-NEITZEL D., ZOCK C., DOERFLER W.;
RL   J. VIROL. 68:379-389(1994).
DR   EMBL; X73487; G313372; -.
DR   PIR; S33938; S33938.
KW   LATE PROTEIN.
SQ   SEQUENCE   497 AA;  56393 MW;  A5BEC571 CRC32;

  Query Match           58.0%;  Score 47;  DB 1;  Length 497;
  Best Local Similarity  66.7%;  Pred. No. 3.12e+00;
  Matches    6;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db    310 ETDPKGRSY 318
          |:|| |:||
Qy      4 EADPTGHSY 12


RESULT   26
ID   XRCC_HUMAN     STANDARD;      PRT;   633 AA.
AC   P18887;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE   DNA-REPAIR PROTEIN XRCC1.
GN   XRCC1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91061722.
RA   THOMPSON L.H., BROOKMAN K.W., JONES N.J., ALLEN S.A., CARRANO A.V.;
RL   MOL. CELL. BIOL. 10:6160-6171(1990).
CC   -!- FUNCTION: CORRECTS DEFECTIVE DNA STRAND-BREAK REPAIR AND SISTER
CC       CHROMATID EXCHANGE FOLLOWING TREATMENT WITH IONIZING RADIATION
CC       AND ALKYLATING AGENTS.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC   -!- SIMILARITY: SOME, TO S.POMBE RAD4/CUT5.
DR   EMBL; M36089; G340397; -.
DR   PIR; A36353; A36353.
DR   MIM; 194360; -.
KW   DNA REPAIR; NUCLEAR PROTEIN.
SQ   SEQUENCE   633 AA;  69525 MW;  D382BC2E CRC32;

  Query Match           58.0%;  Score 47;  DB 1;  Length 633;
  Best Local Similarity  54.5%;  Pred. No. 3.12e+00;
  Matches    6;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db    201 VTASDPAGPSY 211
          |  :||:| ||
Qy      2 VKEADPTGHSY 12


RESULT   27
ID   PA24_BUNMU     STANDARD;      PRT;   147 AA.
AC   P17934;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE   PHOSPHOLIPASE A2, BETA BUNGAROTOXIN A4 CHAIN PRECURSOR (EC 3.1.1.4)
DE   (PHOSPHATIDYLCHOLINE 2-ACYLHYDROLASE).
OS   BUNGARUS MULTICINCTUS (MANY-BANDED KRAIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC   LEPIDOSAURIA; SERPENTES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RX   MEDLINE; 90356417.
RA   DANSE J.M., GARNIER J.M., KEMPF J.;
RL   NUCLEIC ACIDS RES. 18:4610-4610(1990).
CC   -!- FUNCTION: INHIBITS NEUROMUSCULAR TRANSMISSION BY BLOCKING
CC       ACETYLCHOLINE RELEASE FROM THE NERVE TERMINI. ACT PRESYNAPTICALLY.
CC   -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC       2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC       PHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- SUBUNIT: DIMER OF DISSIMILAR CHAINS LINKED BY ONE OR MORE
CC       DISULFIDE BONDS. THE A CHAINS HAVE PHOSPHOLIPASE A2 ACTIVITY AND
CC       THE B CHAINS SHOW HOMOLOGY WITH THE BASIC PROTEASE INHIBITORS.
DR   EMBL; X53408; G62506; -.
DR   PIR; S10982; PSKFA4.
DR   HSSP; P15445; 1PSH.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
KW   HYDROLASE; LIPID DEGRADATION; CALCIUM; MULTIGENE FAMILY; VENOM;
KW   PRESYNAPTIC NEUROTOXIN; SIGNAL.
FT   SIGNAL        1     27
FT   CHAIN        28    147       BETA BUNGAROTOXIN A4 CHAIN.
FT   ACT_SITE     75     75       BY SIMILARITY.
FT   ACT_SITE    119    119       BY SIMILARITY.
FT   DISULFID     42     42       INTERCHAIN (WITH A B CHAIN) (PROBABLE).
FT   DISULFID     54    146       BY SIMILARITY.
FT   DISULFID     56     72       BY SIMILARITY.
FT   DISULFID     71    127       BY SIMILARITY.
FT   DISULFID     78    120       BY SIMILARITY.
FT   DISULFID     88    113       BY SIMILARITY.
FT   DISULFID    106    118       BY SIMILARITY.
FT   CA_BIND      76     76       BY SIMILARITY.
SQ   SEQUENCE   147 AA;  16177 MW;  57165FE9 CRC32;

  Query Match           56.8%;  Score 46;  DB 1;  Length 147;
  Best Local Similarity  33.3%;  Pred. No. 5.40e+00;
  Matches    4;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db     84 NIRDCDPKTQSY 95
          :::: ||  :||
Qy      1 DVKEADPTGHSY 12


RESULT   28
ID   ARA2_ECOLI     STANDARD;      PRT;   583 AA.
AC   P52145;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   ARSENICAL PUMP-DRIVING ATPASE (EC 3.6.1.-).
GN   ARSA.
OS   ESCHERICHIA COLI.
OG   PLASMID R46.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96275894.
```

```
RA   BRUHN D.F., LI J., SILVER S., ROBERTO F., ROSEN B.P.;
RL   FEMS MICROBIOL. LETT. 139:149-153(1996).
CC   -!- FUNCTION: ANION-TRANSPORTING ATPASE. CATALYSES THE EXTRUSION
CC       OF THE OXYANIONS ARSENITE, ANTIMONITE AND ARSENATE. MAINTENANCE
CC       OF A LOW INTRACELLULAR CONCENTRATION OF OXYANION PRODUCES
CC       RESISTANCE TO THE TOXIC AGENTS.
DR   EMBL; U38947; G1061416; -.
KW   PLASMID; ARSENICAL RESISTANCE; ATP-BINDING.
FT   NP_BIND      15     22       ATP (POTENTIAL).
FT   NP_BIND     334    341       ATP (POTENTIAL).
SQ   SEQUENCE   583 AA;  63521 MW;  1071F37E CRC32;

  Query Match           56.8%;  Score 46;  DB 1;  Length 583;
  Best Local Similarity  50.0%;  Pred. No. 5.40e+00;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    445 VMDTAPTGHT 454
          |::: ||||:
Qy      2 VKEADPTGHS 11


RESULT   29
ID   ARA1_ECOLI     STANDARD;      PRT;   583 AA.
AC   P08690;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   ARSENICAL PUMP-DRIVING ATPASE (EC 3.6.1.-).
GN   ARSA.
OS   ESCHERICHIA COLI.
OG   PLASMID R773.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87033737.
RA   CHEN C.-M., MISRA T.K., SILVER S., ROSEN B.P.;
RL   J. BIOL. CHEM. 261:15030-15038(1986).
RN   [2]
RP   REVIEW.
RX   MEDLINE; 91126299.
RA   ROSEN B.P.;
RL   RES. MICROBIOL. 141:336-341(1990).
CC   -!- FUNCTION: ANION-TRANSPORTING ATPASE. CATALYSES THE EXTRUSION
CC       OF THE OXYANIONS ARSENITE, ANTIMONITE AND ARSENATE. MAINTENANCE
CC       OF A LOW INTRACELLULAR CONCENTRATION OF OXYANION PRODUCES
CC       RESISTANCE TO THE TOXIC AGENTS.
DR   EMBL; J02591; G151857; -.
DR   PIR; A25937; A25937.
KW   PLASMID; ARSENICAL RESISTANCE; ATP-BINDING.
FT   NP_BIND      15     22       ATP (POTENTIAL).
FT   NP_BIND     334    341       ATP (POTENTIAL).
SQ   SEQUENCE   583 AA;  63188 MW;  79BAEF0A CRC32;

  Query Match           56.8%;  Score 46;  DB 1;  Length 583;
  Best Local Similarity  50.0%;  Pred. No. 5.40e+00;
  Matches     5;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    445 VMDTAPTGHT 454
          |::: ||||:
Qy      2 VKEADPTGHS 11


RESULT   30
ID   VP57_BDV       STANDARD;      PRT;   503 AA.
AC   P52638;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   57 KD PROTEIN (P57).
OS   BORNA DISEASE VIRUS (BDV).
OC   VIRIDAE; SS-RNA ENVELOPED VIRUSES; NEGATIVE-STRAND; BORNAVIRIDAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V;
RX   MEDLINE; 94240137.
RA   BRIESE T., SCHNEEMANN A., LEWIS A.J., PARK Y.-S., KIM S.,
RA   LUDWIG H., LIPKIN W.I.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 91:4362-4366(1994).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
DR   EMBL; U04608; G516506; -.
KW   GLYCOPROTEIN; TRANSMEMBRANE.
FT   TRANSMEM      5     25       POTENTIAL.
FT   TRANSMEM    274    294       POTENTIAL.
FT   TRANSMEM    468    488       POTENTIAL.
FT   CARBOHYD     63     63       POTENTIAL.
FT   CARBOHYD    109    109       POTENTIAL.
FT   CARBOHYD    139    139       POTENTIAL.
FT   CARBOHYD    192    192       POTENTIAL.
FT   CARBOHYD    196    196       POTENTIAL.
FT   CARBOHYD    202    202       POTENTIAL.
FT   CARBOHYD    221    221       POTENTIAL.
FT   CARBOHYD    230    230       POTENTIAL.
FT   CARBOHYD    235    235       POTENTIAL.
FT   CARBOHYD    321    321       POTENTIAL.
FT   CARBOHYD    328    328       POTENTIAL.
FT   CARBOHYD    388    388       POTENTIAL.
FT   CARBOHYD    438    438       POTENTIAL.
SQ   SEQUENCE   503 AA;  56652 MW;  5B493E19 CRC32;

  Query Match           55.6%;  Score 45;  DB 1;  Length 503;
  Best Local Similarity  55.6%;  Pred. No. 9.26e+00;
  Matches     5;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    416 ETDPINHAY 424
          |:|| |:|
Qy      4 EADPTGHSY 12


RESULT   31
ID   PBP2_NEIGO     STANDARD;      PRT;   581 AA.
AC   P08149;
DT   01-AUG-1988 (REL. 08, CREATED)
DT   01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   PENICILLIN-BINDING PROTEIN 2 (PBP-2).
GN   PENA.
OS   NEISSERIA GONORRHOEAE.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   NEISSERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88156937.
RA   SPRATT B.G.;
RL   NATURE 332:173-176(1988).
CC   -!- CATALYTIC ACTIVITY: SYNTHESIS OF CROSS-LINKED PEPTIDOGLYCAN FROM
CC       THE LIPID INTERMEDIATES. THE ENZYME HAS AN N-TERMINAL PENICILLIN
CC       INSENSITIVE TRANGLYCOSYLASE DOMAIN (FORMATION OF LINEAR GLYCAN
CC       STRANDS) & A CARBOXY-TERMINAL PENICILLIN-SENSITIVE TRANSPEPTIDASE
CC       DOMAIN (CROSS-LINKING OF THE PEPTIDE SUBUNITS).
CC   -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: INNER MEMBRANE.
CC   -!- THIS PROTEIN WAS SEQUENCED IN PENICILLIN-SENSITIVE STRAINS LM306,
CC       AND FA19, AND IN PENICILLIN-RESISTANT STRAINS CDC84-060384,
CC       CDC84-060418 AND CDC77-124615. THE SEQUENCE SHOWN IS THAT OF
CC       STRAIN LM306.
DR   EMBL; M32091; G150279; -.
DR   EMBL; X07468; G44911; -.
DR   EMBL; X07469; G44913; -.
DR   EMBL; X07470; G44915; -.
DR   PIR; S00916; S00916.
KW   INNER MEMBRANE; PEPTIDOGLYCAN SYNTHESIS; CELL DIVISION; CELL WALL;
KW   ANTIBIOTIC RESISTANCE; MULTIFUNCTIONAL ENZYME.
FT   ACT_SITE    310    310       ACYLATED BY PENICILLIN.
FT   VARIANT     346    346       D -> DD (IN CDC84-060418, CDC77-124615,
FT                                AND CDC84-060384).
```

```
FT   VARIANT     504    504       F -> L (IN CDC84-060418, CDC77-124615,
FT                                 AND CDC84-060384).
FT   VARIANT     510    510       A -> V (IN CDC84-060418, CDC77-124615,
FT                                 AND CDC84-060384).
FT   VARIANT     516    516       A -> G (IN CDC84-060418, CDC77-124615,
FT                                 AND CDC84-060384).
FT   VARIANT     541    541       H -> N (IN FA19 AND CDC84-060418).
FT   VARIANT     551    551       P -> S (IN CDC77-124615).
FT   VARIANT     551    551       P -> L (IN CDC84-060384).
FT   VARIANT     552    552       P -> V (IN CDC84-060418).
FT   VARIANT     555    556       KI -> QV (IN CDC84-060418).
FT   VARIANT     566    566       I -> V (IN CDC84-060418).
FT   VARIANT     574    574       A -> NV (IN CDC84-060418).
SQ   SEQUENCE   581 AA;  63650 MW;  D1275110 CRC32;

  Query Match           55.6%;  Score 45;  DB 1;  Length 581;
  Best Local Similarity  45.5%;  Pred. No. 9.26e+00;
  Matches    5;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db    533 VTIDEPTAHGY 543
          |   :||:|:|
Qy      2 VKEADPTGHSY 12


RESULT   32
ID   PBP2_NEIME     STANDARD;      PRT;   581 AA.
AC   P11882;
DT   01-OCT-1989 (REL. 12, CREATED)
DT   01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   PENICILLIN-BINDING PROTEIN 2 (PBP-2).
GN   PENA.
OS   NEISSERIA MENINGITIDIS.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC   NEISSERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SEROGROUP B, STRAIN C311;
RX   MEDLINE; 89345099.
RA   XHANG Q.-Y., SPRATT B.G.;
RL   NUCLEIC ACIDS RES. 17:5383-5383(1989).
CC   -!- CATALYTIC ACTIVITY: SYNTHESIS OF CROSS-LINKED PEPTIDOGLYCAN FROM
CC       THE LIPID INTERMEDIATES..THE ENZYME HAS AN N-TERMINAL PENICILLIN
CC       INSENSITIVE TRANGLYCOSYLASE DOMAIN (FORMATION OF LINEAR GLYCAN
CC       STRANDS) & A CARBOXY-TERMINAL PENICILLIN-SENSITIVE TRANSPEPTIDASE
CC       DOMAIN (CROSS-LINKING OF THE PEPTIDE SUBUNITS).
CC   -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: INNER MEMBRANE.
DR   EMBL; X15276; G45178; -.
DR   PIR; S04857; S04857.
KW   INNER MEMBRANE; PEPTIDOGLYCAN SYNTHESIS; CELL DIVISION; CELL WALL;
KW   ANTIBIOTIC RESISTANCE; MULTIFUNCTIONAL ENZYME.
FT   ACT_SITE    310    310       ACYLATED BY PENICILLIN.
SQ   SEQUENCE   581 AA;  63604 MW;  806E7D60 CRC32;

  Query Match           55.6%;  Score 45;  DB 1;  Length 581;
  Best Local Similarity  45.5%;  Pred. No. 9.26e+00;
  Matches    5;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db    533 VTIDEPTAHGY 543
          |   :||:|:|
Qy      2 VKEADPTGHSY 12


RESULT   33
ID   MYT1_HUMAN     STANDARD;      PRT;   725 AA.
AC   Q01538;
DT   01-JUL-1993 (REL. 26, CREATED)
DT   01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   MYELIN TRANSCRIPTION FACTOR 1 (MYT1) (MYTI) (PROTEOLIPID PROTEIN
DE   BINDING PROTEIN) (PLPB1) (FRAGMENT).
GN   PLPB1 OR MTF1 OR MYTI.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93078764.
RA   KIM J.G., HUDSON L.D.;
RL   MOL. CELL. BIOL. 12:5632-5639(1992).
CC   -!- FUNCTION: BINDS TO THE PROMOTER REGIONS OF PROTEOLIPID PROTEINS
CC       OF THE CENTRAL NERVOUS SYSTEM.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: MOSTLY IN DEVELOPING NERVOUS SYSTEM.
CC   -!- DOMAIN: CONTAINS SIX ZINC FINGERS OF THE C2HC CLASS ARRANGED IN
CC       TWO WIDELY SEPARATED CLUSTERS. THESE TWO DOMAINS OF DNA BINDING
CC       CAN FUNCTION INDEPENDENTLY AND RECOGNIZE THE SAME DNA SEQUENCE.
DR   EMBL; M96980; G189042; -.
DR   PIR; A45033; A45033.
DR   PIR; S27964; S27964.
DR   MIM; 600379; -.
KW   TRANSCRIPTION REGULATION; ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN;
KW   REPEAT.
FT   NON_TER       1      1
FT   ZN_FING      17     43       C2HC-TYPE.
FT   ZN_FING      61     87       C2HC-TYPE.
FT   ZN_FING     402    428       C2HC-TYPE.
FT   ZN_FING     451    472       C2HC-TYPE.
FT   ZN_FING     495    521       C2HC-TYPE.
FT   ZN_FING     548    574       C2HC-TYPE.
SQ   SEQUENCE   725 AA;  79149 MW;  6143A6FA CRC32;

  Query Match           55.6%;  Score 45;  DB 1;  Length 725;
  Best Local Similarity  33.3%;  Pred. No. 9.26e+00;
  Matches    4;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    348 EPEESEPAAHSF 359
          :  |::|::||:
Qy      1 DVKEADPTGHSY 12


RESULT   34
ID   T1R1_ECOLI     STANDARD;      PRT;  1033 AA.
AC   P10486;
DT   01-JUL-1989 (REL. 11, CREATED)
DT   01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   TYPE I RESTRICTION ENZYME ECOR124II R PROTEIN (EC 3.1.21.3).
GN   HSDR OR HSR.
OS   ESCHERICHIA COLI.
OG   PLASMID R124/3.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89178628.
RA   PRICE C., LINGNER J., BICKLE J., FIRMAN T.A., GLOVER S.W.;
RL   J. MOL. BIOL. 205:115-125(1989).
CC   -!- FUNCTION: THE ECOR124/3 I ENZYME RECOGNIZES 5'GAA(N7)RTCG.
CC   -!- FUNCTION: SUBUNIT R IS REQUIRED FOR BOTH NUCLEASE AND ATPASE
CC       ACTIVITIES, BUT NOT FOR MODIFICATION.
CC   -!- SUBUNIT: THE TYPE I RESTRICTION & MODIFICATION SYSTEM IS COMPOSED
CC       OF THREE POLYPEPTIDES R,M AND S.
CC   -!- TYPE I RESTRICTION AND MODIFICATION ENZYMES ARE COMPLEX, MULTI-
CC       FUNCTIONAL SYSTEMS WHICH REQUIRE ATP, S-ADENOSYL METHIONINE AND
CC       MG(2+) AS CO-FACTORS AND, IN ADDITION TO THEIR ENDONUCLEOLYTIC
CC       AND METHYLASE ACTIVITIES, ARE POTENT DNA-DEPENDENT ATPASES.
CC   -!- SIMILARITY: WITH ATPASES.
DR   EMBL; X13145; G41750; -.
DR   PIR; S02168; S02168.
DR   REBASE; RB00748; EcoR124II.
KW   PLASMID; RESTRICTION SYSTEM; DNA-BINDING; ATP-BINDING.
SQ   SEQUENCE   1033 AA;  119656 MW;  9E988CC1 CRC32;
```

```
  Query Match           55.6%;  Score 45;  DB 1;  Length 1033;
  Best Local Similarity  75.0%;  Pred. No. 9.26e+00;
  Matches     6;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db     23 AEPTGDSY 30
          |:||| ||
Qy      5 ADPTGHSY 12


RESULT   35
ID   YHV4_LACHE     STANDARD;      PRT;    22 AA.
AC   P22296;
DT   01-AUG-1991 (REL. 19, CREATED)
DT   01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT   01-AUG-1991 (REL. 19, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL PROTEIN IN HLV 3'REGION (ORF4) (FRAGMENT).
OS   LACTOBACILLUS HELVETICUS.
OC   PROKARYOTA; FIRMICUTES; REGULAR ASPOROGENOUS ROD; LACTOBACILLACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=481;
RX   MEDLINE; 91035244.
RA   JOERGER M.C., KLAENHAMMER T.R.;
RL   J. BACTERIOL. 172:6339-6347(1990).
DR   EMBL; M59360; G149556; -.
DR   PIR; D37145; D37145.
KW   HYPOTHETICAL PROTEIN.
FT   NON_TER       22     22
SQ   SEQUENCE   22 AA;  2484 MW;  DE81C16E CRC32;

  Query Match           54.3%;  Score 44;  DB 1;  Length 22;
  Best Local Similarity  40.0%;  Pred. No. 1.57e+01;
  Matches     4;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db      9 KDGNSTGQKY 18
          |::::||: |
Qy      3 KEADPTGHSY 12


Search completed: Tue Apr  7 08:41:45 1998
Job time : 16 secs.
```

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:42:35 1998; MasPar time 5.37 Seconds
95.819 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-4
Description: (1-12) from 5541104.pep
Perfect Score: 81
Sequence: 1 DVKEADPTGHSY 12

Scoring table: PAM 150
Gap 15

Searched: 195121 seqs, 42852602 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: pir55
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 18.565; Variance 47.138; scale 0.394

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	81	100.0	280	2	JC2358	TOIG of: jc2358 check	8.03e-03
2	74	91.4	234	2	I38667	A;Title: Structure, ch	7.31e-02
3	74	91.4	315	2	I38668	A;Title: Structure, ch	7.31e-02
4	72	88.9	369	2	I38659	A;Title: Structure, ch	1.36e-01
5	71	87.7	319	2	I38660	A;Title: Structure, ch	1.85e-01
6	62	76.5	317	2	JC2359	A;Accession: PH1298.	2.84e+00
7	59	72.8	317	2	I38661	A;Accession: PH1297.	6.86e+00
8	57	70.4	347	2	I38008	TOIG of: i38008 check	1.23e+01
9	53	65.4	314	2	JC2360	A;Title: Structure, ch	3.85e+01
10	52	64.2	417	2	S16582	TOIG of: s16582 check	5.09e+01
11	51	63.0	334	2	E69361	A;Accession: E69361.	6.73e+01
12	50	61.7	314	2	JC2361	A;Title: Human gene MA	8.87e+01
13	50	61.7	381	2	S29560	A;Accession: S29560.	8.87e+01
14	50	61.7	835	2	A49891	TOIG of: a49891 check	8.87e+01
15	49	60.5	417	2	A43929	A;Accession: A39246.	1.17e+02
16	49	60.5	3828	1	MELITNELLY	This is a DE line.	1.17e+02
17	48	59.3	90	2	I46953	A;Cross-references: GB	1.53e+02
18	48	59.3	150	2	S55370	A;Cross-references: EM	1.53e+02
19	48	59.3	269	2	A49334	TOIG of: a49334 check	1.53e+02
20	48	59.3	866	2	C64834	A;Title: The complete	1.53e+02
21	48	59.3	875	2	G69910	A;Status: preliminary;	1.53e+02
22	48	59.3	925	2	A39216	A;Accession: S51030.	1.53e+02
23	47	58.0	197	1	KIPGGU	TOIG of: kipggu check	2.01e+02
24	47	58.0	445	2	A48947	A;Contents: L. d. bulg	2.01e+02
25	47	58.0	497	2	S33938	TOIG of: s33938 check	2.01e+02
26	47	58.0	517	1	S	TOIG of: ajbsqu check	2.01e+02
27	47	58.0	517	1	S	A;Experimental source:	2.01e+02
28	47	58.0	633	2	A36353	A;Accession: A36353.	2.01e+02
29	46	56.8	9	2	PH1299	A;Accession: PH1299.	2.62e+02
30	46	56.8	9	2	PH1300	A;Accession: PH1300.	2.62e+02
31	46	56.8	147	1	PSKFA4	C;Keywords: calcium; c	2.62e+02
32	46	56.8	382	2	S54702	A;Accession: S54702.	2.62e+02
33	46	56.8	439	2	S25483	TOIG of: s25483 check	2.62e+02
34	46	56.8	551	2	S65289	TOIG of: s65289 check	2.62e+02
35	46	56.8	551	2	S66740	TOIG of: s66740 check	2.62e+02
36	46	56.8	583	1	A25937	F;117-121/Region: nucl	2.62e+02
37	45	55.6	46	2	C69745	A;Experimental source:	3.42e+02
38	45	55.6	125	2	D49923	TOIG of: d49923 check	3.42e+02
39	45	55.6	581	2	S00916	TOIG of: s00916 check	3.42e+02
40	45	55.6	581	2	S04857	TOIG of: s04857 check	3.42e+02
41	45	55.6	582	2	S49090	A;Cross-references: EM	3.42e+02
42	45	55.6	644	2	G64938	A;Title: The complete	3.42e+02
43	45	55.6	725	2	A45033	TOIG of: a45033 check	3.42e+02
44	45	55.6	1033	2	S02168	TOIG of: s02168 check	3.42e+02
45	44	54.3	22	2	D37145	A;Accession: D37145.	4.44e+02
46	44	54.3	61	2	S60796	A;Cross-references: EM	4.44e+02
47	44	54.3	222	2	S39681	A;Title: Bacillus subt	4.44e+02
48	44	54.3	222	2	E69722	A;Status: preliminary;	4.44e+02
49	44	54.3	261	2	S52899	TOIG of: s52899 check	4.44e+02
50	44	54.3	377	2	S52537	A;Accession: S52537.	4.44e+02
51	44	54.3	438	2	S55631	TOIG of: s55631 check	4.44e+02
52	44	54.3	467	2	S75150	A;Accession: S75150.	4.44e+02
53	44	54.3	488	2	A53107	F;25-103/Domain: cytoc	4.44e+02
54	44	54.3	488	2	S55874	TOIG of: s55874 check	4.44e+02
55	44	54.3	539	2	D36904	A;Accession: D36904.	4.44e+02
56	44	54.3	608	2	E69513	A;Accession: E69513.	4.44e+02
57	44	54.3	629	2	S54567	TOIG of: s54567 check	4.44e+02
58	44	54.3	678	1	323	TOIG of: ajzrql check	4.44e+02
59	44	54.3	789	2	S46631	A;Accession: S46631.	4.44e+02
60	44	54.3	827	1	A31642	F;17-350/Domain: gelso	4.44e+02
61	43	53.1	177	1	R5HSL5	A;Accession: I28949.	5.76e+02
62	43	53.1	229	2	S66577	A;Cross-references: EM	5.76e+02
63	43	53.1	290	2	S64312	TOIG of: s64312 check	5.76e+02
64	43	53.1	301	2	A37766	C;Superfamily: cellula	5.76e+02
65	43	53.1	326	2	S44753	TOIG of: s44753 check	5.76e+02
66	43	53.1	363	2	I64150	A;Authors: Geoghagen,	5.76e+02
67	43	53.1	370	2	S49008	A;Accession: B56556.	5.76e+02
68	43	53.1	390	2	D69531	A;Accession: D69531.	5.76e+02
69	43	53.1	559	2	B36307	TOIG of: b36307 check	5.76e+02
70	43	53.1	564	2	G00041	TOIG of: g00041 check	5.76e+02
71	43	53.1	621	2	S37664	TOIG of: s37664 check	5.76e+02
72	43	53.1	669	2	I38029	A;Accession: I38029.	5.76e+02
73	43	53.1	686	2	A44842	TOIG of: a44842 check	5.76e+02
74	43	53.1	688	2	B42161	TOIG of: b42161 check	5.76e+02
75	43	53.1	690	2	A42161	TOIG of: a42161 check	5.76e+02
76	43	53.1	690	2	S07103	A;Accession: S07103.	5.76e+02
77	43	53.1	762	2	I59329	F;286-408/Domain: cAMP	5.76e+02
78	43	53.1	796	2	A32434	A;Accession: A32434.	5.76e+02
79	43	53.1	878	2	S44543	TOIG of: s44543 check	5.76e+02
80	43	53.1	2504	2	A57788	F;492-773/Domain: [acy	5.76e+02
81	43	53.1	2504	2	B57788	F;492-773/Domain: [acy	5.76e+02
82	43	53.1	2509	2	G01880	F;2123-2198/Domain: ac	5.76e+02
83	43	53.1	3224	2	S58884	A;Title: A giant nucle	5.76e+02
84	43	53.1	3224	2	A57545	A;Cross-references: GB	5.76e+02
85	43	53.1	3396	1	A42551	C;Date: 30-Sep-1993 #s	5.76e+02
86	42	51.9	156	1	GNVQWA	TOIG of: gnvqwa check	7.44e+02
87	42	51.9	156	2	E69882	A;Status: preliminary;	7.44e+02
88	42	51.9	283	2	E69626	A;Status: preliminary;	7.44e+02
89	42	51.9	342	2	S63404	TOIG of: s63404 check	7.44e+02
90	42	51.9	348	1	DEEBOT	C;Superfamily: dihydro	7.44e+02
91	42	51.9	379	2	A42421	TOIG of: a42421 check	7.44e+02
92	42	51.9	445	2	C69735	A;Experimental source:	7.44e+02
93	42	51.9	446	2	MEGSVLTLVL	This is a DE line.	7.44e+02
94	42	51.9	644	1	DUCTIVE	TOIG of: rgaswa check	7.44e+02
95	42	51.9	661	2	S49901	TOIG of: s49901 check	7.44e+02
96	42	51.9	700	2	S38928	TOIG of: s38928 check	7.44e+02

```
 97     42   51.9    773  2  S46011      A;Accession: S46620.    7.44e+02
 98     42   51.9    799  2  A34729      TOIG of: a34729  check  7.44e+02
 99     42   51.9    960  2  UCH         TOIG of: b47093  check  7.44e+02
100     42   51.9   1132  2  S37932      A;Cross-references: EM  7.44e+02
```

ALIGNMENTS

```
RESULT    1
ID   JC2358     STANDARD;     PRT;   280 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: jc2358  check: 467  from: 1  to: 280.
XX
CC     TOIG of: jc2358  check: 467  from: 1  to: 280
CC
CC    >P1;JC2358
CC    tumor-associated antigen , MAGE-1 - human
CC    C;Species: Homo sapiens (man)
CC    C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Mar-19
  96
CC    C;Accession: JC2358
CC    R;Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
CC    Biochem. Biophys. Res. Commun. 202, 549-555, 1994
CC    A;Title: Cloning and analysis of MAGE-1-related genes.
CC    A;Reference number: JC2358
CC    A;Accession: JC2358
CC    A;Molecule type: mRNA
CC    A;Residues: 1-280 <DIN>
CC    A;Experimental source: melanoma cell line DM150
CC    C;Genetics:
CC    A;Gene: MAGE
CC    F;161-169/Region: HLA-A1 binding #status predicted
SQ   SEQUENCE   280 AA; 30932 MW; 426797 CN;

  Query Match           100.0%;  Score 81;  DB 2;  Length 280;
  Best Local Similarity 100.0%;  Pred. No. 8.03e-03;
  Matches    12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     158 DVKEADPTGHSY 169
           ||||||||||||
Qy       1 DVKEADPTGHSY 12


RESULT    2
ID   I38667     STANDARD;     PRT;   234 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: Structure, chromosomal localization, and expression of 12 genes o
  f the MAGE family.
XX
CC    A;Title: Structure, chromosomal localization, and expression of 12 genes
  of the MAGE family.
CC    A;Reference number: I38659; MUID:95012457
CC    A;Accession: I38667
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-234 <RES>
CC    A;Cross-references: EMBL:U10693; NID:g533525; PID:g533526
CC    C;Genetics:
CC    A;Gene: GDB:MAGE8
CC    A;Cross-references: GDB:331123
CC    A;Map position: Xq28-Xq28
CC    A;Introns: #status absent
SQ   SEQUENCE   234 AA; 25197 MW; 296950 CN;

  Query Match           91.4%;  Score 74;  DB 2;  Length 234;
  Best Local Similarity 83.3%;  Pred. No. 7.31e-02;
  Matches    10;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db     168 DVKEVDPAGHSY 179
           |||| ||:||||
Qy       1 DVKEADPTGHSY 12


RESULT    3
ID   I38668     STANDARD;     PRT;   315 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: Structure, chromosomal localization, and expression of 12 genes o
  f the MAGE family.
XX
CC    A;Title: Structure, chromosomal localization, and expression of 12 genes
  of the MAGE family.
CC    A;Reference number: I38659; MUID:95012457
CC    A;Accession: I38668
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-315 <RES>
CC    A;Cross-references: EMBL:U10694; NID:g533527; PID:g533528
CC    C;Genetics:
CC    A;Gene: GDB:MAGE9
CC    A;Cross-references: GDB:331125
CC    A;Map position: Xp21.3-Xp21.3
CC    A;Introns: #status absent
SQ   SEQUENCE   315 AA; 35088 MW; 553046 CN;

  Query Match           91.4%;  Score 74;  DB 2;  Length 315;
  Best Local Similarity 83.3%;  Pred. No. 7.31e-02;
  Matches    10;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db     164 DVKEVDPAGHSY 175
           |||| ||:||||
Qy       1 DVKEADPTGHSY 12


RESULT    4
ID   I38659     STANDARD;     PRT;   369 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: Structure, chromosomal localization, and expression of 12 genes o
  f the MAGE family.
XX
CC    A;Title: Structure, chromosomal localization, and expression of 12 genes
  of the MAGE family.
CC    A;Reference number: I38659; MUID:95012457
CC    A;Accession: I38659
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-369 <RES>
CC    A;Cross-references: EMBL:U10685; NID:g533510; PID:g533511
CC    C;Genetics:
CC    A;Gene: GDB:MAGE10
CC    A;Cross-references: GDB:331126
CC    A;Map position: Xq28-Xq28
CC    A;Introns: #status absent
SQ   SEQUENCE   369 AA; 40766 MW; 735605 CN;

  Query Match           88.9%;  Score 72;  DB 2;  Length 369;
  Best Local Similarity 83.3%;  Pred. No. 1.36e-01;
  Matches    10;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db     190 DVKEVDPTGHSF 201
           |||| ||||||:
```

```
Qy      1 DVKEADPTGHSY 12


RESULT    5
ID   I38660     STANDARD;     PRT;   319 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: Structure, chromosomal localization, and expression of 12 genes o
  f the MAGE family.
XX
CC    A;Title: Structure, chromosomal localization, and expression of 12 genes
  of the MAGE family.
CC    A;Reference number: I38659; MUID:95012457
CC    A;Accession: I38660
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-319 <RES>
CC    A;Cross-references: EMBL:U10686; NID:g533512; PID:g533513
CC    C;Genetics:
CC    A;Gene: GDB:MAGE11
CC    A;Cross-references: GDB:331128
CC    A;Map position: Xq28-Xq28
CC    A;Introns: #status absent
SQ   SEQUENCE   319 AA; 35536 MW; 548364 CN;

  Query Match           87.7%;  Score 71;  DB 2;  Length 319;
  Best Local Similarity  83.3%;  Pred. No. 1.85e-01;
  Matches    10;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    168 DVKEVDPTSHSY 179
          |||| |||:|||
Qy      1 DVKEADPTGHSY 12


RESULT    6
ID   JC2359     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: PH1298.
XX
CC    A;Accession: PH1298
CC    A;Molecule type: DNA
CC    A;Residues: 169-177 <TRA>
CC    R;Fenton, R.G.
CC    submitted to the EMBL Data Library, June 1994
CC    A;Reference number: G07128
CC    A;Accession: G01446
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: mRNA
CC    A;Residues: 1-317 <FEN>
CC    A;Cross-references: EMBL:U10340; NID:g499123; PID:g499124
CC    C;Genetics:
CC    A;Gene: MAGE-X2
CC    F;169-177/Region: HLA-A1 binding #status predicted
SQ   SEQUENCE   317 AA; 34928 MW; 530585 CN;

  Query Match           76.5%;  Score 62;  DB 2;  Length 317;
  Best Local Similarity  66.7%;  Pred. No. 2.84e+00;
  Matches     8;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    166 DVKEVDPTSNTY 177
          |||| |||:::|
Qy      1 DVKEADPTGHSY 12


RESULT    7
ID   I38661     STANDARD;     PRT;   317 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: PH1297.
XX
CC    A;Accession: PH1297
CC    A;Molecule type: DNA
CC    A;Residues: 169-177 <TRA>
CC    C;Genetics:
CC    A;Gene: GDB:MAGE4
CC    A;Cross-references: GDB:331119
CC    A;Map position: Xq28-Xq28
CC    A;Introns: #status absent
SQ   SEQUENCE   317 AA; 34899 MW; 528124 CN;

  Query Match           72.8%;  Score 59;  DB 2;  Length 317;
  Best Local Similarity  58.3%;  Pred. No. 6.86e+00;
  Matches     7;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    166 DVKEVDPASNTY 177
          |||| ||:::|
Qy      1 DVKEADPTGHSY 12


RESULT    8
ID   I38008     STANDARD;     PRT;   347 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: i38008  check: 8233  from: 1  to: 347.
XX
CC    TOIG of: i38008  check: 8233  from: 1  to: 347
CC
CC    >P1;I38008
CC    MAGE-Xp protein - human
CC    C;Species: Homo sapiens (man)
CC    C;Date: 01-Mar-1996 #sequence_revision 01-Mar-1996 #text_change 06-Sep-19
  96
CC    C;Accession: I38008; S52167
CC    R;Muscatelli, F.; Walker, A.P.; De Plaen, E.; Stafford, A.N.; Monaco, A.P
.
CC    Proc. Natl. Acad. Sci. U.S.A. 92, 4987-4991, 1995
CC    A;Title: Isolation and characterization of a MAGE gene family in the Xp21
  .3 region.
CC    A;Reference number: I38008; MUID:95281581
CC    A;Accession: I38008
CC    A;Status: preliminary
CC    A;Molecule type: mRNA
CC    A;Residues: 1-347 <RES>
CC    A;Cross-references: EMBL:X82539; NID:g608992; PID:g608993
CC    C;Genetics:
CC    A;Gene: GDB:MAGEL1-LSB
CC    A;Cross-references: GDB:635712
CC    A;Map position: Xp21.3-Xp21.3
SQ   SEQUENCE   347 AA; 39152 MW; 610197 CN;

  Query Match           70.4%;  Score 57;  DB 2;  Length 347;
  Best Local Similarity  50.0%;  Pred. No. 1.23e+01;
  Matches     6;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db    164 DLKEDNPSSHTY 175
          |:|| :|::|:|
Qy      1 DVKEADPTGHSY 12


RESULT    9
ID   JC2360     STANDARD;     PRT;   314 AA.
XX
AC   xxxxxx
```

```
XX
DT   01-JAN-1900
XX
DE   A;Title: Structure, chromosomal localization, and expression of 12 genes o
  f the MAGE family.
XX
CC    A;Title: Structure, chromosomal localization, and expression of 12 genes
  of the MAGE family.
CC    A;Reference number: I38659; MUID:95012457
CC    A;Accession: I38665
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-314 <RES>
CC    A;Cross-references: EMBL:U10691; NID:g533522; PID:g533523
CC    R;Fenton, R.G.
CC    submitted to the EMBL Data Library, June 1994
CC    A;Reference number: G07126
CC    A;Accession: G01445
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: mRNA
CC    A;Residues: 1-314 <FEN>
CC    A;Cross-references: EMBL:U10339; NID:g499121; PID:g499122
CC    C;Genetics:
CC    A;Gene: GDB:MAGE6
CC    A;Cross-references: GDB:331121
CC    A;Map position: Xq28-Xq28
CC    A;Introns: #status absent
CC    F;168-176/Region: HLA-A1 binding #status predicted
SQ   SEQUENCE   314 AA; 34891 MW; 534374 CN;

  Query Match           65.4%;  Score 53;  DB 2;  Length 314;
  Best Local Similarity  50.0%;  Pred. No. 3.85e+01;
  Matches    6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db    165 ELMEVDPIGHVY 176
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT   10
ID   S16582     STANDARD;     PRT;   417 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s16582  check: 7441  from: 1  to: 417.
XX
CC     TOIG of: s16582  check: 7441  from: 1  to: 417
CC
CC    >P1;S16582
CC    fructose-bisphosphatase (EC 3.1.3.11) precursor, chloroplast - Arabidopsi
  s thaliana
CC    C;Species: Arabidopsis thaliana (mouse-ear cress)
CC    C;Date: 21-Nov-1993 #sequence_revision 12-May-1995 #text_change 08-Sep-19
  97
CC    C;Accession: S16582
CC    R;Horsnell, P.R.; Raines, C.A.
CC    Plant Mol. Biol. 17, 185-186, 1991
CC    A;Title: Nucleotide sequence of a cDNA clone encoding chloroplast fructos
  e-1,6-bisphosphatase from Arabidopsis thaliana.
CC    A;Reference number: S16582; MUID:91329733
CC    A;Accession: S16582
CC    A;Molecule type: mRNA
CC    A;Residues: 1-417 <HOR>
CC    A;Cross-references: EMBL:X58148; NID:g11241; PID:g11242
CC    A;Experimental source: clone AFBP1
CC    C;Genetics:
CC    A;Genome: nuclear
CC    C;Superfamily: fructose-bisphosphatase
CC    C;Keywords: Calvin cycle; chloroplast; gluconeogenesis; phosphoric monoes
  ter hydrolase
CC    F;1-58/Domain: transit peptide (chloroplast) #status predicted <TNP>
CC    F;59-417/Product: fructose-bisphosphatase #status predicted <MAT>
SQ   SEQUENCE   417 AA; 45177 MW; 917183 CN;

  Query Match           64.2%;  Score 52;  DB 2;  Length 417;
  Best Local Similarity  50.0%;  Pred. No. 5.09e+01;
  Matches    6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    314 DLKDPGPTGKPY 325
          |:|:: ||| :|
Qy      1 DVKEADPTGHSY 12


RESULT   11
ID   E69361     STANDARD;     PRT;   334 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: E69361.
XX
CC    A;Accession: E69361
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-334 <KLE>
CC    A;Cross-references: GB:AE000782; TIGR:AF0893
SQ   SEQUENCE   334 AA; 38033 MW; 575945 CN;

  Query Match           63.0%;  Score 51;  DB 2;  Length 334;
  Best Local Similarity  41.7%;  Pred. No. 6.73e+01;
  Matches    5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    314 EVKDNEPTGTVF 325
          :||: :|||  :
Qy      1 DVKEADPTGHSY 12


RESULT   12
ID   JC2361     STANDARD;     PRT;   314 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: Human gene MAGE-3 codes for an antigen recognized on a melanoma b
  y autologous cytolytic T lymphocytes.
XX
CC    A;Title: Human gene MAGE-3 codes for an antigen recognized on a melanoma
  by autologous cytolytic T lymphocytes.
CC    A;Reference number: I38438; MUID:94157413
CC    A;Accession: I38438
CC    A;Status: preliminary; translated from GB/EMBL/DDBJ
CC    A;Molecule type: DNA
CC    A;Residues: 1-314 <RES>
CC    A;Cross-references: EMBL:U03735; NID:g468825; PID:g468826
CC    C;Genetics:
CC    A;Gene: MAGE-3
CC    F;168-176/Region: HLA-A1 binding #status predicted
SQ   SEQUENCE   314 AA; 34747 MW; 538982 CN;

  Query Match           61.7%;  Score 50;  DB 2;  Length 314;
  Best Local Similarity  50.0%;  Pred. No. 8.87e+01;
  Matches    6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db    165 ELMEVDPIGHLY 176
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT   13
ID   S29560     STANDARD;     PRT;   381 AA.
XX
```

```
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S29560.
XX
CC    A;Accession: S29560
CC    A;Molecule type: mRNA
CC    A;Residues: 1-381 <CAR>
CC    A;Cross-references: EMBL:X68826; NID:g20716; PID:g20717
CC    C;Superfamily: fructose-bisphosphatase
CC    C;Keywords: phosphoric monoester hydrolase
SQ   SEQUENCE   381 AA; 41821 MW; 781009 CN;

  Query Match           61.7%;  Score 50;  DB 2;  Length 381;
  Best Local Similarity  50.0%;  Pred. No. 8.87e+01;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    278 DLKEPGPSGKPY 289
          |:||: |:| :|
Qy      1 DVKEADPTGHSY 12


RESULT   14
ID   A49891     STANDARD;     PRT;   835 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: a49891  check: 6404  from: 1  to: 835.
XX
CC     TOIG of: a49891  check: 6404  from: 1  to: 835
CC
CC    >P1;A49891
CC    outer membrane protein FasD precursor - Escherichia coli
CC    C;Species: Escherichia coli
CC    C;Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 14-Nov-19
  97
CC    C;Accession: A49891
CC    R;Schifferli, D.M.; Alrutz, M.A.
CC    J. Bacteriol. 176, 1099-1110, 1994
CC    A;Title: Permissive linker insertion sites in the outer membrane protein
  of 987P fimbriae of Escherichia coli.
CC    A;Reference number: A49891
CC    A;Accession: A49891
CC    A;Status: preliminary
CC    A;Molecule type: DNA
CC    A;Residues: 1-835 <SCH>
CC    A;Cross-references: GB:L22659; NID:g437334; PID:g437336
CC    C;Genetics:
CC    A;Gene: fasD
CC    C;Superfamily: hypothetical protein b0532
CC    C;Keywords: membrane protein
SQ   SEQUENCE   835 AA; 92353 MW; 3824200 CN;

  Query Match           61.7%;  Score 50;  DB 2;  Length 835;
  Best Local Similarity  50.0%;  Pred. No. 8.87e+01;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db    324 NIKEADGSEHSF 335
          ::|||| : ||:
Qy      1 DVKEADPTGHSY 12


RESULT   15
ID   A43929     STANDARD;     PRT;   417 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: A39246.
XX
CC    A;Accession: A39246
CC    A;Molecule type: mRNA
CC    A;Residues: 1-417 <RE2>
CC    A;Cross-references: GB:M27717; NID:g179933; PID:g179934
CC    R;Goldstein, S.M.; Kaempfer, C.E.; Kealey, J.T.; Wintroub, B.U.
CC    J. Clin. Invest. 83, 1630-1636, 1989
CC    A;Title: Human mast cell carboxypeptidase. Purification and characterizat
  ion.
CC    A;Reference number: A45759
CC    A;Accession: A45759
CC    A;Molecule type: protein
CC    A;Residues: 110-137 <GOL>
CC    C;Genetics:
CC    A;Gene: GDB:CPA3
CC    A;Cross-references: GDB:125231; OMIM:114851
CC    A;Map position: 3q21.3-3q25
CC    A;Introns: 23/2; 48/3; 90/2; 124/3; 158/3; 192/3; 229/3; 260/1; 327/3; 35
  6/1
CC    C;Superfamily: carboxypeptidase
CC    C;Keywords: hydrolase; metallo-carboxypeptidase; metalloprotein; protein
  digestion; zinc; zymogen
CC    F;1-15/Domain: signal sequence #status predicted <SIG>
CC    F;16-109/Domain: activation peptide #status predicted <ACT>
CC    F;110-417/Product: carboxypeptidase A, mast cell #status predicted <MAT>
CC    F;176,179,304/Binding site: zinc (His, Glu, His) #status predicted
CC    F;245-268/Disulfide bonds: #status predicted
CC    F;356,378/Active site: Tyr, Glu #status predicted
SQ   SEQUENCE   417 AA; 48700 MW; 936226 CN;

  Query Match           60.5%;  Score 49;  DB 2;  Length 417;
  Best Local Similarity  66.7%;  Pred. No. 1.17e+02;
  Matches     8;  Conservative    0;  Mismatches   4;  Indels   0;  Gaps   0;

Db    105 DVKEDIPGRHSY 116
          ||||  |  |||
Qy      1 DVKEADPTGHSY 12


RESULT   16
ID   MELITNELLYKTYKQKPVGVEEPVYDQAGNPLFGERGAIHPQSTLKLPHKRGERNVPTSLASLPKRGDCR
   STANDARD;     PRT;   3828 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   This is a DE line.
XX
SQ   SEQUENCE   3828 AA; 429949 MW; 77459508 CN;

  Query Match           60.5%;  Score 49;  DB 1;  Length 3828;
  Best Local Similarity  58.3%;  Pred. No. 1.17e+02;
  Matches     7;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db   1062 DVVKADPGGQGY 1073
          ||  ||| |::|
Qy      1 DVKEADPTGHSY 12


RESULT   17
ID   I46953     STANDARD;     PRT;   90 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Cross-references: GB:S65218; NID:g410551; PID:g410552.
XX
CC    A;Cross-references: GB:S65218; NID:g410551; PID:g410552
SQ   SEQUENCE   90 AA; 10981 MW; 46468 CN;

  Query Match           59.3%;  Score 48;  DB 2;  Length 90;
```

```
  Best Local Similarity  41.7%;  Pred. No. 1.53e+02;
  Matches     5;  Conservative    5;  Mismatches   2;  Indels   0;  Gaps   0;

Db      22 EVMIIDPSGNTY 33
           :|:  ||:|::|
Qy       1 DVKEADPTGHSY 12


RESULT   18
ID   S55370     STANDARD;     PRT;   150 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Cross-references: EMBL:Z49199; NID:g854176; PID:g854177.
XX
CC    A;Cross-references: EMBL:Z49199; NID:g854176; PID:g854177
SQ   SEQUENCE   150 AA; 17143 MW; 115920 CN;

  Query Match           59.3%;  Score 48;  DB 2;  Length 150;
  Best Local Similarity  50.0%;  Pred. No. 1.53e+02;
  Matches     6;  Conservative    2;  Mismatches   4;  Indels   0;  Gaps   0;

Db      11 DVKDIDPEGKKF 22
           |||: || |  :
Qy       1 DVKEADPTGHSY 12


RESULT   19
ID   A49334     STANDARD;     PRT;   269 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: a49334  check: 9237  from: 1  to: 269.
XX
CC     TOIG of: a49334  check: 9237  from: 1  to: 269
CC
CC    >P1;A49334
CC    Ras homolog Rad - human
CC    N;Alternate names: Ras associated with diabetes (Rad)
CC    C;Species: Homo sapiens (man)
CC    C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-Oct-19
  94
CC    C;Accession: A49334
CC    R;Reynet, C.; Kahn, C.R.
CC    Science 262, 1441-1444, 1993
CC    A;Title: Rad: a member of the Ras family overexpressed in muscle of type
  II diabetic humans.
CC    A;Reference number: A49334
CC    A;Accession: A49334
CC    A;Status: preliminary
CC    A;Molecule type: mRNA
CC    A;Residues: 1-269 <REY>
CC    A;Cross-references: GB:L24564
CC    C;Keywords: alternative initiators
SQ   SEQUENCE   269 AA; 29262 MW; 334168 CN;

  Query Match           59.3%;  Score 48;  DB 2;  Length 269;
  Best Local Similarity  55.6%;  Pred. No. 1.53e+02;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db      80 EAEAAGHTY 88
           ||:::||:|
Qy       4 EADPTGHSY 12


RESULT   20
ID   C64834     STANDARD;     PRT;   866 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Title: The complete genome sequence of Escherichia coli K-12.
XX
CC    A;Title: The complete genome sequence of Escherichia coli K-12.
CC    A;Reference number: A64720; MUID:97426617
CC    A;Accession: C64834
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-866 <BLAT>
CC    A;Cross-references: GB:AE000196; GB:U00096; NID:g1787169; PID:g1787172; U
  WGP:b0940
CC    A;Experimental source: strain K-12, substrain MG1655
CC    C;Superfamily: hypothetical protein b0532
CC    C;Keywords: membrane protein
SQ   SEQUENCE   866 AA; 95241 MW; 3791425 CN;

  Query Match           59.3%;  Score 48;  DB 2;  Length 866;
  Best Local Similarity  50.0%;  Pred. No. 1.53e+02;
  Matches     6;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     334 EIKEADGSVNSY 345
           ::|||| : :||
Qy       1 DVKEADPTGHSY 12


RESULT   21
ID   G69910     STANDARD;     PRT;   875 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Status: preliminary; nucleic acid sequence not shown; translation not sh
  own.
XX
CC    A;Status: preliminary; nucleic acid sequence not shown; translation not s
  hown
CC    A;Molecule type: DNA
CC    A;Residues: 1-875 <KUN>
CC    A;Experimental source: strain 168
CC    C;Genetics:
CC    A;Gene: yomG
SQ   SEQUENCE   875 AA; 98884 MW; 4045045 CN;

  Query Match           59.3%;  Score 48;  DB 2;  Length 875;
  Best Local Similarity  45.5%;  Pred. No. 1.53e+02;
  Matches     5;  Conservative    4;  Mismatches   2;  Indels   0;  Gaps   0;

Db     231 INSVNPTGQSY 241
           ::  :|||:||
Qy       2 VKEADPTGHSY 12


RESULT   22
ID   A39216     STANDARD;     PRT;   925 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S51030.
XX
CC    A;Accession: S51030
CC    A;Status: preliminary
CC    A;Molecule type: mRNA
CC    A;Residues: 1-80 <BEL>
CC    C;Genetics:
CC    A;Gene: GDB:PDNP1; M6S1; NPPS
CC    A;Cross-references: GDB:132615; OMIM:173335
CC    A;Map position: 6q22-6q23
```

```
CC    C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology
CC    C;Keywords: glycoprotein; phosphoric diester hydrolase; transmembrane pro
  tein
CC    F;77-97/Domain: transmembrane #status predicted <TMM>
CC    F;104-144/Domain: somatomedin B homology <SBH1>
CC    F;145-188/Domain: somatomedin B homology <SBH2>
CC    F;179,285,341,477,578,585,643,700,731,748/Binding site: carbohydrate (Asn
  ) (covalent) #status predicted
CC    F;254/Active site: Thr (covalent substrate-binding) #status predicted
SQ   SEQUENCE   925 AA; 104924 MW; 4589090 CN;

  Query Match           59.3%;  Score 48;  DB 2;  Length 925;
  Best Local Similarity  66.7%;  Pred. No. 1.53e+02;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db    374 EPDSSGHSY 382
          |:|::|||| 
Qy      4 EADPTGHSY 12


RESULT   23
ID   KIPGGU     STANDARD;     PRT;   197 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: kipggu  check: 2414  from: 1  to: 197.
XX
CC     TOIG of: kipggu  check: 2414  from: 1  to: 197
CC
CC    >P1;KIPGGU
CC    guanylate kinase (EC 2.7.4.8) - pig
CC    C;Species: Sus scrofa domestica (domestic pig)
CC    C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 23-Aug-19
  97
CC    C;Accession: S23776; S32545
CC    R;Zschocke, P.D.; Schiltz, E.; Schulz, G.E.
CC    submitted to the Protein Sequence Database, September 1992
CC    A;Reference number: S23776
CC    A;Accession: S23776
CC    A;Molecule type: protein
CC    A;Residues: 1-197 <ZSC>
CC    R;Zschocke, P.D.; Schiltz, E.; Schulz, G.E.
CC    Eur. J. Biochem. 213, 263-269, 1993
CC    A;Title: Purification and sequence determination of guanylate kinase from
  pig brain.
CC    A;Reference number: S32545
CC    A;Accession: S32545
CC    A;Status: preliminary
CC    A;Molecule type: protein
CC    A;Residues: 1-197 <ZS2>
CC    C;Superfamily: guanylate kinase; guanylate kinase homology
CC    C;Keywords: acetylated amino end; ATP; magnesium; monomer; P-loop; phosph
  otransferase
CC    F;3-188/Domain: guanylate kinase homology <GKI>
CC    F;10-17/Region: nucleotide-binding motif A (P-loop)
CC    F;35-82/Region: GMP binding #status predicted
CC    F;1/Modified site: acetylated amino end (Gly) #status experimental
CC    F;16/Binding site: ATP (Lys) #status predicted
SQ   SEQUENCE   197 AA; 21789 MW; 183102 CN;

  Query Match           58.0%;  Score 47;  DB 1;  Length 197;
  Best Local Similarity  54.5%;  Pred. No. 2.01e+02;
  Matches     6;  Conservative    4;  Mismatches   1;  Indels   0;  Gaps   0;

Db    187 EIKKAQATGHS 197
          ::| |::|||| 
Qy      1 DVKEADPTGHS 11


RESULT   24
ID   A48947     STANDARD;     PRT;   445 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Contents: L. d. bulgaricus.
XX
CC    A;Contents: L. d. bulgaricus
CC    A;Accession: A48947
CC    A;Status: preliminary; not compared with conceptual translation
CC    A;Molecule type: DNA
CC    A;Residues: 1-445 <ISH>
CC    A;Experimental source: ATCC 11842
CC    A;Note: sequence extracted from NCBI backbone (NCBIP:117804)
CC    C;Superfamily: glutamate--ammonia ligase
CC    C;Keywords: ligase
SQ   SEQUENCE   445 AA; 50133 MW; 1031509 CN;

  Query Match           58.0%;  Score 47;  DB 2;  Length 445;
  Best Local Similarity  33.3%;  Pred. No. 2.01e+02;
  Matches     4;  Conservative    7;  Mismatches   1;  Indels   0;  Gaps   0;

Db    335 EMRSTDPTANPY 346
          ::: :|||:::| 
Qy      1 DVKEADPTGHSY 12


RESULT   25
ID   S33938     STANDARD;     PRT;   497 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s33938  check: 2182  from: 1  to: 497.
XX
CC     TOIG of: s33938  check: 2182  from: 1  to: 497
CC
CC    >P1;S33938
CC    penton protein (III) - human adenovirus 12
CC    C;Species: Mastadenovirus h12 (human adenovirus 12)
CC    C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Sep-19
  97
CC    C;Accession: S33938
CC    R;Sprengel, J.
CC    submitted to the EMBL Data Library, June 1993
CC    A;Reference number: S33928
CC    A;Accession: S33938
CC    A;Status: preliminary
CC    A;Molecule type: DNA
CC    A;Residues: 1-497 <SPR>
CC    A;Cross-references: EMBL:X73487; NID:g313361; PID:g313372
CC    C;Superfamily: adenovirus penton protein
SQ   SEQUENCE   497 AA; 56393 MW; 1345686 CN;

  Query Match           58.0%;  Score 47;  DB 2;  Length 497;
  Best Local Similarity  66.7%;  Pred. No. 2.01e+02;
  Matches     6;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db    310 ETDPKGRSY 318
          |:|| |:|| 
Qy      4 EADPTGHSY 12


RESULT   26
ID   S BOTH MN2+ AND MG2+ IONS FOR HIGH ACTIVITY.     STANDARD;     PRT;   539
  AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: ajbsqu  check: 2505  from: 1  to: 444.
```

```
XX
CC     TOIG of: ajbsqu  check: 2505  from: 1  to: 444
CC
CC    >P1;AJBSQU
CC    glutamate--ammonia ligase (EC 6.3.1.2) - Bacillus cereus
CC    N;Alternate names: glutamine synthetase
CC    C;Species: Bacillus cereus
CC    C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-Sep-19
  97
CC    C;Accession: JU0075; A38060
CC    R;Nakano, Y.; Kato, C.; Tanaka, E.; Kimura, K.; Horikoshi, K.
CC    J. Biochem. 106, 209-215, 1989
CC    A;Title: Nucleotide sequence of the glutamine synthetase gene (glnA) and
  its upstream region from Bacillus cereus.
CC    A;Reference number: A91912; MUID:90036764
CC    A;Accession: JU0075
CC    A;Molecule type: DNA
CC    A;Residues: 1-444 <NAK>
CC    A;Cross-references: GB:D00513; NID:g216271; PID:d1000857; PID:g216273
CC    A;Accession: A38060
CC    A;Molecule type: protein
CC    A;Residues: 1-20 <NA2>
CC    C;Comment: This enzyme catalyzes the formation of glutamine from ammonia
  and glutamic acid in the presence of ATP. It requir
SQ   SEQUENCE   539 AA; 60400 MW; 1449741 CN;

  Query Match           58.0%;  Score 47;  DB 1;  Length 517;
  Best Local Similarity  33.3%;  Pred. No. 2.01e+02;
  Matches     4;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    406 EVRSVDPAANPY 417
          :|:  ||::::|
Qy      1 DVKEADPTGHSY 12


RESULT   27
ID   S BOTH MN2+ AND MG2+ IONS FOR HIGH ACTIVITY.     STANDARD;     PRT;   539
  AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Experimental source: strain 168.
XX
CC    A;Experimental source: strain 168
CC    C;Comment: This enzyme catalyzes the formation of glutamine from ammonia
  and glutamic acid in the presence of ATP. It requir
SQ   SEQUENCE   539 AA; 60460 MW; 1472580 CN;

  Query Match           58.0%;  Score 47;  DB 1;  Length 517;
  Best Local Similarity  33.3%;  Pred. No. 2.01e+02;
  Matches     4;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    406 EVRSVDPAANPY 417
          :|:  ||::::|
Qy      1 DVKEADPTGHSY 12


RESULT   28
ID   A36353     STANDARD;     PRT;   633 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: A36353.
XX
CC    A;Accession: A36353
CC    A;Status: preliminary
CC    A;Molecule type: mRNA
CC    A;Residues: 1-633 <THO>
CC    A;Cross-references: GB:M36089
CC    C;Genetics:
CC    A;Gene: GDB:XRCC1; RCC
CC    A;Cross-references: GDB:120737; OMIM:194360
CC    A;Map position: 19q13.2-19q13.2
SQ   SEQUENCE   633 AA; 69525 MW; 2018971 CN;

  Query Match           58.0%;  Score 47;  DB 2;  Length 633;
  Best Local Similarity  54.5%;  Pred. No. 2.01e+02;
  Matches     6;  Conservative    2;  Mismatches   3;  Indels   0;  Gaps   0;

Db    201 VTASDPAGPSY 211
          |  :||:| ||
Qy      2 VKEADPTGHSY 12


RESULT   29
ID   PH1299     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: PH1299.
XX
CC    A;Accession: PH1299
CC    A;Molecule type: DNA
CC    A;Residues: 1-9 <TRA>
SQ   SEQUENCE   9 AA; 997 MW; 590 CN;

  Query Match           56.8%;  Score 46;  DB 2;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 2.62e+02;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTSNTY 9
          |||||:::|
Qy      4 EADPTGHSY 12


RESULT   30
ID   PH1300     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: PH1300.
XX
CC    A;Accession: PH1300
CC    A;Molecule type: DNA
CC    A;Residues: 1-9 <TRA>
SQ   SEQUENCE   9 AA; 997 MW; 590 CN;

  Query Match           56.8%;  Score 46;  DB 2;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 2.62e+02;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTSNTY 9
          |||||:::|
Qy      4 EADPTGHSY 12


RESULT   31
ID   PSKFA4     STANDARD;     PRT;   147 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   C;Keywords: calcium; carboxylic ester hydrolase; lipid degradation; metall
  oprotein; toxin; venom.
XX
CC    C;Keywords: calcium; carboxylic ester hydrolase; lipid degradation; metal
  loprotein; toxin; venom
```

```
CC    F;1-27/Domain: signal sequence #status predicted <SIG>
CC    F;28-147/Product: phospholipase A2 isoform A4 #status predicted <MAT>
CC    F;54-146,56-72,71-127,78-120,88-113,106-118/Disulfide bonds: #status pred
  icted
CC    F;75,121/Active site: His, Asp #status predicted
SQ   SEQUENCE   147 AA; 16177 MW; 97291 CN;

  Query Match           56.8%;  Score 46;  DB 1;  Length 147;
  Best Local Similarity  33.3%;  Pred. No. 2.62e+02;
  Matches     4;  Conservative    5;  Mismatches   3;  Indels   0;  Gaps   0;

Db     84 NIRDCDPKTQSY 95
          :::: ||  :||
Qy      1 DVKEADPTGHSY 12


RESULT   32
ID   S54702     STANDARD;     PRT;   382 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   A;Accession: S54702.
XX
CC    A;Accession: S54702
CC    A;Status: preliminary
CC    A;Molecule type: DNA
CC    A;Residues: 1-382 <WHI>
CC    A;Cross-references: EMBL:L27667; NID:g443685; PID:g443686
SQ   SEQUENCE   382 AA; 42532 MW; 685630 CN;

  Query Match           56.8%;  Score 46;  DB 2;  Length 382;
  Best Local Similarity  25.0%;  Pred. No. 2.62e+02;
  Matches     3;  Conservative    6;  Mismatches   3;  Indels   0;  Gaps   0;

Db    370 EITDDDPAGRRF 381
          :: : ||:|: :
Qy      1 DVKEADPTGHSY 12


RESULT   33
ID   S25483     STANDARD;     PRT;   439 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s25483  check: 3624  from: 1  to: 439.
XX
CC     TOIG of: s25483  check: 3624  from: 1  to: 439
CC
CC    >P1;S25483
CC    ribulose-bisphosphate carboxylase activase (clone JQ4) - common tobacco
CC    N;Alternate names: rubisco activase
CC    C;Species: Nicotiana tabacum (common tobacco)
CC    C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Sep-19
  97
CC    C;Accession: S25483
CC    R;Rodermel, S.; Qian, J.
CC    submitted to the EMBL Data Library, August 1992
CC    A;Description: Characterization of three Rubisco activase cDNAs from toba
  cco.
CC    A;Reference number: S25482
CC    A;Accession: S25483
CC    A;Molecule type: DNA
CC    A;Residues: 1-439 <ROD>
CC    A;Cross-references: EMBL:Z14980; NID:g19989; PID:g19990
SQ   SEQUENCE   439 AA; 48343 MW; 941507 CN;

  Query Match           56.8%;  Score 46;  DB 2;  Length 439;
  Best Local Similarity  41.7%;  Pred. No. 2.62e+02;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db     60 EEKDADPKKQTY 71
          : |:|||  ::|
Qy      1 DVKEADPTGHSY 12


RESULT   34
ID   S65289     STANDARD;     PRT;   551 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s65289  check: 3373  from: 1  to: 551.
XX
CC     TOIG of: s65289  check: 3373  from: 1  to: 551
CC
CC    >P1;S65289
CC    hypothetical protein YPL258c - yeast (Saccharomyces cerevisiae)
CC    N;Alternate names: hypothetical protein P0701
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 14-Nov-19
  97
CC    C;Accession: S65289
CC    R;Messenguy, F.; Dubois, E.; Vierendeels, F.; Scherens, B.
CC    submitted to the Protein Sequence Database, May 1996
CC    A;Reference number: S64935
CC    A;Accession: S65289
CC    A;Molecule type: DNA
CC    A;Residues: 1-551 <MES>
CC    A;Cross-references: EMBL:Z73614; NID:g1370531; PID:e246962; PID:g1370532;
   MIPS:YPL258c
CC    A;Experimental source: strain S288C (AB972)
CC    C;Genetics:
CC    A;Map position: 16L
SQ   SEQUENCE   551 AA; 61334 MW; 1568767 CN;

  Query Match           56.8%;  Score 46;  DB 2;  Length 551;
  Best Local Similarity  45.5%;  Pred. No. 2.62e+02;
  Matches     5;  Conservative    1;  Mismatches   5;  Indels   0;  Gaps   0;

Db    285 VKDNGPINHVY 295
          ||:  |  | |
Qy      2 VKEADPTGHSY 12


RESULT   35
ID   S66740     STANDARD;     PRT;   551 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   TOIG of: s66740  check: 1641  from: 1  to: 551.
XX
CC     TOIG of: s66740  check: 1641  from: 1  to: 551
CC
CC    >P1;S66740
CC    probable transcription factor YOL055c - yeast (Saccharomyces cerevisiae)
CC    N;Alternate names: protein O1239
CC    C;Species: Saccharomyces cerevisiae
CC    C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Nov-19
  97
CC    C;Accession: S66740; S66747; S59294; S61724
CC    R;Ansorge, W.; Benes, V.; Rechmann, S.; Schwager, C.; Teodoru, C.; Voss,
  H.; Wiemann, S.
CC    submitted to the Protein Sequence Database, July 1996
CC    A;Reference number: S66723
CC    A;Accession: S66740
CC    A;Molecule type: DNA
CC    A;Residues: 1-551 <ANS>
CC    A;Cross-references: EMBL:Z74797; NID:g1419864; PID:e251864; PID:g1419865;
   MIPS:YOL055c
```

```
CC    A;Experimental source: strain S288C
CC    R;Feldmann, H.; Mannhaupt, G.; Vetter, I.
CC    submitted to the Protein Sequence Database, July 1996
CC    A;Reference number: S66743
CC    A;Accession: S66747
CC    A;Molecule type: DNA
CC    A;Residues: 1-551 <FEL>
CC    A;Cross-references: EMBL:Z74797; NID:g1419864; PID:e251864; PID:g1419865;
   MIPS:YOL055c
CC    A;Experimental source: strain S288C
CC    R;Mannhaupt, G.; Vetter, I.; Schwarzlose, C.; Mitzel, S.; Feldmann, H.
CC    submitted to the EMBL Data Library, August 1995
CC    A;Reference number: S59285
CC    A;Accession: S59294
CC    A;Molecule type: DNA
CC    A;Residues: 1-543 <FEW>
CC    A;Cross-references: EMBL:X91067; NID:g984177; PID:g984187
CC    R;Mannhaupt, G.; Vetter, I.; Schwarzlose, C.; Mitzel, S.; Feldmann, H.
CC    Yeast 12, 67-76, 1996
CC    A;Title: Analysis of a 26 kb region on the left arm of yeast chromosome X
  V.
CC    A;Reference number: S61715
CC    A;Accession: S61724
CC    A;Status: nucleic acid sequence not shown; translation not shown
CC    A;Molecule type: DNA
CC    A;Residues: 1-543 <MAN>
CC    A;Cross-references: EMBL:X91067; NID:g984177; PID:g984187
CC    A;Note: the nucleotide sequence was submitted to the EMBL Data Library, A
  ugust 1995
CC    C;Genetics:
CC    A;Map position: 15L
SQ   SEQUENCE   551 AA; 61269 MW; 1606714 CN;

  Query Match           56.8%;  Score 46;  DB 2;  Length 551;
  Best Local Similarity  45.5%;  Pred. No. 2.62e+02;
  Matches     5;  Conservative    1;  Mismatches   5;  Indels   0;  Gaps   0;

Db    285 VKDNGPINHVY 295
          ||:  |  | |
Qy      2 VKEADPTGHSY 12


Search completed: Tue Apr  7 08:42:46 1998
Job time : 11 secs.
```

****

MPSRCH (TM)

****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:43:04 1998; MasPar time 2.70 Seconds
75.061 Million cell updates/sec

Tabular output not generated.

Title: >US-08-190-411A-4
Description: (1-12) from 5541104.pep
Perfect Score: 81
Sequence: 1 DVKEADPTGHSY 12

Scoring table: PAM 150
Gap 15

Searched: 111725 seqs, 16919825 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: a-geneseq30
1:a-geneseq1

Statistics: Mean 21.386; Variance 7.814; scale 2.737

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	81	100.0	38	1	R80620	Immunogenic peptide of	2.05e-34
2	81	100.0	335	1	R70909	Human melanoma antigen	2.05e-34
3	61	75.3	35	1	R49224	HLA-A1 MAGE 1 antigen	1.12e-17
4	61	75.3	35	1	R50281	MAGE-1 nonapeptide.	1.12e-17
5	61	75.3	35	1	W00897	Human melanoma MAGE1 t	1.12e-17
6	61	75.3	35	1	R78824	MAGE-1 cytotoxic T lym	1.12e-17
7	61	75.3	35	1	R47330	HLA-A1 MAGE 1 antigen	1.12e-17
8	61	75.3	35	1	R65135	MAGE 1 immunogenic pep	1.12e-17
9	61	75.3	35	1	R99343	MAGE-1 nonapeptide.	1.12e-17
10	61	75.3	35	1	R65112	MAGE 1 immunogenic pep	1.12e-17
11	61	75.3	35	1	R83932	MHC class I restricted	1.12e-17
12	61	75.3	35	1	R63675	Synthetic peptide deri	1.12e-17
13	61	75.3	35	1	R75954	Melanoma antigen (MAGE	1.12e-17
14	61	75.3	35	1	R29769	Antigen E peptide.	1.12e-17
15	61	75.3	35	1	R82988	P815 antigenic peptide	1.12e-17
16	61	75.3	35	1	R90692	Human leukocyte antige	1.12e-17
17	59	72.8	35	1	R99342	HLA binding nonapeptid	3.99e-16
18	58	71.6	35	1	R99339	HLA binding nonapeptid	2.33e-15
19	55	67.9	35	1	R99337	HLA binding nonapeptid	4.16e-13
20	55	67.9	35	1	R99340	HLA binding nonapeptid	4.16e-13
21	54	66.7	35	1	R99338	HLA binding nonapeptid	2.26e-12
22	51	63.0	36	1	W13255	MAGE-3 TRAP HLA-B44 mo	3.22e-10
23	50	61.7	117	1	W16333	Baboon MAGE-3 homologu	1.61e-09
24	50	61.7	122	1	W16332	Baboon MAGE-3 homologu	1.61e-09
25	50	61.7	125	1	W16335	Baboon MAGE-3 homologu	1.61e-09
26	50	61.7	152	1	W16330	Human MAGE-3 tumour an	1.61e-09
27	50	61.7	159	1	W16331	Baboon MAGE-3 homologu	1.61e-09
28	50	61.7	160	1	W16334	Baboon MAGE-3 homologu	1.61e-09
29	49	60.5	35	1	R99341	HLA binding nonapeptid	7.89e-09
30	48	59.3	334	1	W13869	Rad protein.	3.78e-08
31	48	59.3	334	1	R45431	Diabetogene rad: A typ	3.78e-08
32	48	59.3	951	1	R79148	Human insulin receptor	3.78e-08
33	47	58.0	471	1	P93191	Peptide with glutamine	1.77e-07
34	46	56.8	35	1	R99350	MAGE-6 nonapeptide.	8.04e-07
35	46	56.8	35	1	R99349	MAGE-5/MAGE-51 nonapep	8.04e-07
36	46	56.8	35	1	R50288	MAGE-51 nonapeptide.	8.04e-07
37	46	56.8	35	1	R50289	MAGE-6 nonapeptide.	8.04e-07
38	46	56.8	35	1	R50287	MAGE-5 nonapeptide.	8.04e-07
39	46	56.8	1025	1	R87511	Human c-mer protooncog	8.04e-07
40	45	55.6	372	1	R60653	pstS variant.	3.57e-06
41	45	55.6	529	1	R98604	Borna disease virus p5	3.57e-06
42	44	54.3	103	1	P80309	Polypeptide with regio	1.54e-05
43	44	54.3	853	1	R05049	Human villin.	1.54e-05
44	43	53.1	35	1	R50284	MAGE-3 nonapeptide.	6.47e-05
45	43	53.1	35	1	R65118	MAGE 3 immunogenic pep	6.47e-05
46	43	53.1	35	1	R49222	HLA-A1 MAGE 3 antigen	6.47e-05
47	43	53.1	35	1	R99346	MAGE-3 nonapeptide.	6.47e-05
48	43	53.1	35	1	R83931	MHC class I restricted	6.47e-05
49	43	53.1	35	1	R75942	Melanoma antigen (MAGE	6.47e-05
50	43	53.1	37	1	R73851	Antigen fragment 167,	6.47e-05
51	43	53.1	367	1	R77361	Cysteine proteinase fr	6.47e-05
52	43	53.1	695	1	R86408	Human matrix metallopr	6.47e-05
53	43	53.1	838	1	R10047	abaA gene of Aspergill	6.47e-05
54	43	53.1	3422	1	R43662	DEN1-S275/90 (ECACC V9	6.47e-05
55	42	51.9	35	1	R99348	MAGE-41 nonapeptide.	2.64e-04
56	42	51.9	35	1	R50286	MAGE-41 nonapeptide.	2.64e-04
57	42	51.9	182	1	R13587	PLRV viral protein.	2.64e-04
58	42	51.9	224	1	R93137	Mouse guanylate kinase	2.64e-04
59	42	51.9	275	1	W09404	Human hepatoma-derived	2.64e-04
60	42	51.9	418	1	W18668	Poliovirus receptor (4	2.64e-04
61	42	51.9	418	1	R07130	H20B receptor.	2.64e-04
62	42	51.9	442	1	R07131	H20A receptor.	2.64e-04
63	42	51.9	443	1	W18669	Poliovirus receptor (4	2.64e-04
64	41	50.6	345	1	R88011	Mature Pseudomonas glu	1.04e-03
65	41	50.6	369	1	R28887	pRi8 fusion protein.	1.04e-03
66	41	50.6	448	1	P50775	Parasporium rhizobium	1.04e-03
67	41	50.6	485	1	R08331	Hybrid murine IL-7 rec	1.04e-03
68	41	50.6	497	1	R88133	Aureobasidin sensitivi	1.04e-03
69	41	50.6	497	1	R67692	C. albicans caaur1 gen	1.04e-03
70	41	50.6	498	1	R67696	C. albicans caaur1 gen	1.04e-03
71	41	50.6	499	1	W10882	Comamonas acidovorans	1.04e-03
72	41	50.6	556	1	R28888	Ri paraneoplastic anti	1.04e-03
73	41	50.6	748	1	R10055	Cephalosporin C amidas	1.04e-03
74	41	50.6	915	1	R65159	Potassium ion channel	1.04e-03
75	41	50.6	996	1	R72458	Porphyromonas gingival	1.04e-03
76	41	50.6	1113	1	R96028	P. gingivalis haemaggl	1.04e-03
77	41	50.6	1384	1	R96032	P. gingivalis hagD hae	1.04e-03
78	41	50.6	1713	1	R96033	P. gingivalis haemaggl	1.04e-03
79	41	50.6	1730	1	R70188	Arg-gingipain-2 prepol	1.04e-03
80	41	50.6	1732	1	W24786	PrtR antigenic protein	1.04e-03
81	41	50.6	1758	1	R96029	P. gingivalis porphypa	1.04e-03
82	41	50.6	1758	1	W24787	PrtK antigenic protein	1.04e-03
83	41	50.6	3191	1	R38889	Sequence encoded by OR	1.04e-03
84	40	49.4	440	1	W25659	Human cadherin-6.	3.98e-03
85	40	49.4	440	1	W13131	Partial human cadherin	3.98e-03
86	40	49.4	456	1	R27290	Murine osteogenic prot	3.98e-03
87	40	49.4	456	1	R46743	Mouse osteogenic prote	3.98e-03
88	40	49.4	456	1	R85758	mOP-1.	3.98e-03
89	40	49.4	456	1	R57972	Murine OP-1.	3.98e-03
90	40	49.4	456	1	R50199	Murine OP-1.	3.98e-03
91	40	49.4	456	1	R54936	Osteogenic protein mOP	3.98e-03
92	40	49.4	456	1	R50237	Murine OP-1.	3.98e-03
93	40	49.4	456	1	R47291	mOP1.	3.98e-03
94	40	49.4	456	1	R47251	mOP1.	3.98e-03
95	40	49.4	456	1	R51645	mOP1-PP prepro form mo	3.98e-03
96	40	49.4	456	1	W00237	Mouse osteogenic prote	3.98e-03

97	40	49.4	549	1	R71976	Pertussis A.	3.98e-03
98	40	49.4	1444	1	R71703	Collagen alpha 1 (II)	3.98e-03
99	40	49.4	1444	1	R59751	Type II collagen.	3.98e-03
100	40	49.4	1859	1	R79478	Mouse LTBP-2.	3.98e-03

ALIGNMENTS

```
RESULT    1
ID   R80620 standard; Protein; 12 AA.
AC   R80620;
DT   28-FEB-1996  (first entry)
DE   Immunogenic peptide of tumour rejection antigen (MAGE-1).
KW   Tumour rejection antigen; MAGE-1; monoclonal antibody; MAb;
KW   diagnosis; immunoassay; cancer; immunogen; antisera.
OS   Homo sapiens.
PN   WO9520974-A1.
PD   10-AUG-1995.
PF   05-JAN-1995; U00095.
PR   01-FEB-1994; US-190411.
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
PA   (SLOK ) MEMORIAL SLOAN-KETTERING CANCER CENT.
PI   Boon-falleur T,  Chen Y,  Garin-chesa P,  Old LJ,  Rettig WJ;
PI   Stockert E,  Van der bruggen P;
DR   WPI; 95-283606/37.
PT   New monoclonal antibody binding specifically to MAGE-1 - useful for
PT   diagnosis and monitoring of cancer, also new hybridomas, recombinant
PT   MAGE-1 and immunogenic peptide(s)
PS   Claim 12; Page 20; 33pp; English.
CC   A monoclonal antibody directed against the tumour rejection antigen
CC   (MAGE-1) can be used to detect MAGE-1 in samples by standard
CC   immunoassay methods for diagnosis and monitoring of cancer etc. The
CC   monoclonal antibody is designated MA454 and is produced by the
CC   hybridoma deposited as ATCC HBB11540.  The monoclonal antibody is
CC   specific for MAGE-1, having no reactivity for MAGE-2 or MAGE-3.
CC   Peptide fragments of MAGE-1 (See R80618-20) may be useful as
CC   immunogens for production of the monoclonal antibody and antisera.
SQ   Sequence   12 AA;

  Query Match          100.0%;  Score 81;  DB 1;  Length 38;
  Best Local Similarity 100.0%;  Pred. No. 2.05e-34;
  Matches   12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 DVKEADPTGHSY 38
           ||||||||||||
Qy      1 DVKEADPTGHSY 12


RESULT    2
ID   R70909 standard; Protein; 309 AA.
AC   R70909;
DT   09-OCT-1995  (first entry)
DE   Human melanoma antigen MAGE-1.
KW   Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
KW   HLA-restricted cytotoxic T-lymphocyte activity.
OS   Homo sapiens.
PN   WO9504542-A.
PD   16-FEB-1995.
PF   02-AUG-1994; U08721.
PR   06-AUG-1993; US-103623.
PA   (CYTE-) CYTEL CORP.
PI   Fikes JD,  Livingston BD,  Sette AD,  Sidney JC;
DR   WPI; 95-090681/12.
DR   N-PSDB; Q85435.
PT   Human melanoma antigen, MAGE-1, peptide(s) - useful for
PT   stimulating immune response against melanoma
PS   Example 1; Fig 1; 59pp; English.
CC   Q85435 encodes R70909 human melanoma antigen MAGE-1, it was used
CC   to produce the C-terminal MAGE-1 peptides described in R70915 to
CC   R70969. These peptides are useful for defining epitopes that
CC   engender a HLA-restricted cytotoxic lymphocyte activity against
CC   MAGE-1 antigens. Compsns. containing these peptides can be
CC   administered, as a vaccine to patients susceptable to MAGE
CC   associated tumours, e.g. melanomas.
SQ   Sequence   309 AA;

  Query Match          100.0%;  Score 81;  DB 1;  Length 335;
  Best Local Similarity 100.0%;  Pred. No. 2.05e-34;
  Matches   12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    184 DVKEADPTGHSY 195
           ||||||||||||
Qy      1 DVKEADPTGHSY 12


RESULT    3
ID   R49224 standard; Protein; 9 AA.
AC   R49224;
DT   31-AUG-1994  (first entry)
DE   HLA-A1 MAGE 1 antigen peptide fragment 958.01.
KW   Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW   immune response; viral infection; cancer; prostate cancer; lymphoma;
KW   hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS   Synthetic.
PN   WO9403205-A.
PD   17-FEB-1994.
PF   06-AUG-1993; U07421.
PR   07-AUG-1992; US-926666.
PR   05-MAR-1993; US-027746.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Grey HM,  Kubo RT,  Sette A;
DR   WPI; 94-065403/08.
PT   Peptide which specifically binds selected MHC allele - used to
PT   induce an immune response for treatment or prevention of viral
PT   infection or cancer, or for diagnosis
PS   Eaxmple 16; Page 116; 150pp; English.
CC   The sequences given in R47304-33 and R49201-44 are immunogenic
CC   peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC   These peptides may be used in the composition of the invention.
CC   These peptides are capable of binding selected MHC molecules and
CC   inducing an immune response.  They can be used to treat and/or
CC   prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC   hepatitis or AIDS.  They can also be used to produce antibodies for
CC   use as diagnostic or therapeutic agents.  The peptides can also be
CC   used as diagnostic agents.
SQ   Sequence   9 AA;

  Query Match          75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPTGHSY 35
           |||||||||
Qy      4 EADPTGHSY 12


RESULT    4
ID   R50281 standard; Protein; 9 AA.
AC   R50281;
DT   26-SEP-1994  (first entry)
DE   MAGE-1 nonapeptide.
KW   MAGE; nonapeptide; cancer; melanoma; breast cancer; HLA;
KW   histocompatability; human leucocyte antigen; probe; treatment;
KW   therapy; vaccine.
OS   Synthetic.
PN   WO9405304-A.
PD   17-MAR-1994.
PF   30-AUG-1993; U08157.
PR   31-AUG-1992; US-938334.
PR   26-MAR-1993; US-037230.
PR   07-JUN-1993; US-073103.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-falleur T,  De Plaen E,  Lurquin C,  Traversari C;
PI   Van Derbruggen P;
DR   WPI; 94-100844/12.
```

```
DR   N-PSDB; Q44751.
PT   New nona:peptide derived from tumour rejection antigen precursor
PT   - presented by HLA-A1 cancer cells, for use in diagnosis or
PT   therapy of esp. melanoma and breast cancer.
PS   Disclosure; Page 19; 33pp; English.
CC   An isolated nonapeptide having the amino acid sequence Glu-Val-Asp-
CC   Pro-Ile-Gly-His-Leu-Tyr is derived from the tumour rejection antigen
CC   precursor encoded by the MAGE-3 gene and presented by HLA-A1.  The
CC   nonapeptide can be used in a vaccine to treat a cancerous condition
CC   involving HLA-A1 subtype cancerous cells.  The nucleic acid encoding
CC   the nonapeptide can be used as a probe to identify tumour cells.
CC   This sequence is homologous to the peptide described and is encoded
CC   by the MAGE-1 gene.
SQ   Sequence   9 AA;

  Query Match           75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      27 EADPTGHSY 35
           |||||||||
Qy       4 EADPTGHSY 12


RESULT    5
ID   W00897 standard; Peptide; 9 AA.
AC   W00897;
DT   23-MAY-1997  (first entry)
DE   Human melanoma MAGE1 tumour associated antigen p161-169.
KW   Adeno-associated virus; vector; liposome; transfection;
KW   dendritic cell; melanoma; MAGE1; adoptive immunotherapy;
KW   tumour associated antigen.
OS   Homo sapiens.
PN   WO9703703-A1.
PD   06-FEB-1997.
PF   19-JUL-1996; U12012.
PR   21-JUL-1995; US-001312.
PR   01-NOV-1995; US-007184.
PR   01-DEC-1995; US-566286.
PA   (RHON ) RHONE POULENC RORER PHARM INC.
PI   Lebkowski JS,  Philip R;
DR   WPI; 97-145208/13.
PT   Adeno-associated virus:liposome complexes for transfecting dendritic
PT   cells - for inducing immune response, useful for treating e.g.
PT   neoplasia or infections
PS   Example 5; Page 58; 134pp; English.
CC   Tumour associated antigens (W13660-61, W00878-903) can be loaded
CC   into dendritic cells and used to induce antitumour immunity.
CC   Alternatively, the dendritic cells are transfected with adeno
CC   associated virus plasmid DNA (which includes DNA encoding the
CC   tumour associated antigen) complexed with cationic liposomes.  The
CC   amtigen loaded or transfected dendritic cells can be used to
CC   generate tumour antigen-specific cytotoxic T lymphocytes for use in
CC   adoptive immunotherapy in a patient having the corresponding
CC   tumour.  A suitable antigen comprises amino acids 161-169 (W00897)
CC   of human melanoma MAGE1.
SQ   Sequence   9 AA;

  Query Match           75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      27 EADPTGHSY 35
           |||||||||
Qy       4 EADPTGHSY 12


RESULT    6
ID   R78824 standard; peptide; 9 AA.
AC   R78824;
DT   26-MAR-1996  (first entry)
DE   MAGE-1 cytotoxic T lymphocyte epitope.
KW   MAGE-1; cytotoxic T; CTL; epitope; helper T; HTL; lymphocyte;
KW   cell; viruses; parasites; tumours; antigens; disease prevention;
KW   treatment.
OS   Homo sapiens.
PN   WO9522317-A1.
PD   24-AUG-1995.
PF   16-FEB-1995; U02121.
PR   16-FEB-1994; US-197484.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Chesnut RW,  Grey H,  Sette AD,  Vitiello MA;
DR   WPI; 95-302545/39.
PT   Compsn. inducing cytotoxic T lymphocyte response to pref. viral,
PT   bacterial, parasitic or tumour antigens - useful in the treatment
PT   and prevention of diseases associated with the antigen e.g.
PT   hepatitis B
PS   Disclosure; Page 17; 109pp; English.
CC   A compsn. which induces a cytotoxic T lymphocyte (CTL) response to
CC   an antigen (Ag) in a mammal comprises, a CTL Ag response inducing
CC   peptide (i.e. R78824-R78853) and a lipid conjugated helper T cell
CC   inducing peptide. The compsn. induces a CTL response to bacterial,
CC   viral or tumour Ags, and is therefore useful in the treatment and
CC   prevention of diseases associated with the Ag.
SQ   Sequence   9 AA;

  Query Match           75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      27 EADPTGHSY 35
           |||||||||
Qy       4 EADPTGHSY 12


RESULT    7
ID   R47330 standard; Protein; 9 AA.
AC   R47330;
DT   31-AUG-1994  (first entry)
DE   HLA-A1 MAGE 1 antigen peptide fragment 161-169.
KW   Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW   immune response; viral infection; cancer; prostate cancer; lymphoma;
KW   hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS   Synthetic.
PN   WO9403205-A.
PD   17-FEB-1994.
PF   06-AUG-1993; U07421.
PR   07-AUG-1992; US-926666.
PR   05-MAR-1993; US-027746.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Grey HM,  Kubo RT,  Sette A;
DR   WPI; 94-065403/08.
PT   Peptide which specifically binds selected MHC allele - used to
PT   induce an immune response for treatment or prevention of viral
PT   infection or cancer, or for diagnosis
PS   Eaxmple 8; Page 52; 150pp; English.
CC   The sequences given in R47304-33 and R49201-44 are immunogenic
CC   peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC   These peptides may be used in the composition of the invention.
CC   These peptides are capable of binding selected MHC molecules and
CC   inducing an immune response.  They can be used to treat and/or
CC   prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC   hepatitis or AIDS.  They can also be used to produce antibodies for
CC   use as diagnostic or therapeutic agents.  The peptides can also be
CC   used as diagnostic agents.
SQ   Sequence   9 AA;

  Query Match           75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      27 EADPTGHSY 35
           |||||||||
Qy       4 EADPTGHSY 12
```

```
RESULT    8
ID   R65135 standard; peptide; 9 AA.
AC   R65135;
DT   09-OCT-1995  (first entry)
DE   MAGE 1 immunogenic peptide A01.
KW   MAGE 1; immunogenic peptide A01; cytotoxic C cells;
KW   in vitro activation; cancer; AIDS; bacterial infections; malaria;
KW   fungal infections; tuberculosis; hepatitis.
OS   Homo sapiens.
PN   WO9504817-A.
PD   16-FEB-1995.
PF   01-AUG-1994; U08672.
PR   06-AUG-1993; US-103401.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Kubo R,  Serra H,  Tsai V,  Wentworth P;
DR   WPI; 95-090895/12.
PT   In vitro activation of cytotoxic T cells for selected killing of
PT   target cells  - for treating e.g. cancer, AIDS, hepatitis etc.by
PT   incubating them with antigen presenting cells loaded with
PT   appropriate immunogenic peptide
PS   Example 3; Page 38; 53pp; English.
CC   R65109-R65145 are immunogenic peptides, they are used in a new
CC   method for the in vitro activation of cytotoxic T cells (CTC).
CC   This is achieved by incubating the CTCs with antigen presenting
CC   cells loaded with an appropriate immunogenic peptide (e.g. one
CC   of the above peptides). By selecting the peptides used the
CC   following diseases and infections can be treated; cancer, AIDS,
CC   hepatitis, other viral and bacterial infections, malaria and
CC   tuberculosis.
SQ   Sequence   9 AA;

  Query Match           75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPTGHSY 35
           |||||||||
Qy      4 EADPTGHSY 12


RESULT    9
ID   R99343 standard; Protein; 9 AA.
AC   R99343;
DT   22-APR-1997  (first entry)
DE   MAGE-1 nonapeptide.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Homo sapiens.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
DR   N-PSDB; T35408.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis  -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Example 4; Fig 4; 41pp; English.
CC   R99343-R99350 represent MAGE nonapeptides, based on the tumour rejection
CC   antigen region of the full length MAGE sequences.  These peptides were
CC   used to design the nonapeptides of the invention (see R99337-R99342),
CC   which bind to a HLA molecule on a cell, and provoke lysis by cytolytic T
CC   cells (CTLs) specific for a complex of the HLA molecule and nonapeptide.
CC   The nonapeptides can be used diagnostically to identify tumours
CC   expressing a particular HLA molecule, or to identify cancer cells.  The
CC   peptides can also be used therapeutically, to induce a CTL response to
CC   tumours (where the peptides are optionally coupled to tumour-specific
CC   antibodies), or to induce a response by CTLs that are otherwise inactive.
CC   The peptide sequences may also be used to expand specific CTLs in vitro
CC   for later return to the patient, such as for treating melanoma.  Tumour
CC   cells can be identified by using DNA encoding the nonapeptides as probes.
CC   Non-human cells transformed with the HLA-A1 gene and a DNA sequence
CC   encoding one of the peptides, can be used to generate CTLs, or to detect
CC   the presence of CTLs in human samples.  The non-human transformed cells,
CC   when polytransformed, are universal effector cells, and can be used in
CC   vaccines, or for treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match           75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPTGHSY 35
           |||||||||
Qy      4 EADPTGHSY 12


RESULT   10
ID   R65112 standard; peptide; 9 AA.
AC   R65112;
DT   06-OCT-1995  (first entry)
DE   MAGE 1 immunogenic peptide 161-169.
KW   MAGE 1; immunogenic peptide 161-169; cytotoxic C cells;
KW   in vitro activation; cancer; AIDS; bacterial infections; malaria;
KW   fungal infections; tuberculosis; hepatitis.
OS   Homo sapiens.
PN   WO9504817-A.
PD   16-FEB-1995.
PF   01-AUG-1994; U08672.
PR   06-AUG-1993; US-103401.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Kubo R,  Serra H,  Tsai V,  Wentworth P;
DR   WPI; 95-090895/12.
PT   In vitro activation of cytotoxic T cells for selected killing of
PT   target cells  - for treating e.g. cancer, AIDS, hepatitis etc.by
PT   incubating them with antigen presenting cells loaded with
PT   appropriate immunogenic peptide
PS   Example 3; Page 35; 53pp; English.
CC   R65109-R65145 are immunogenic peptides, they are used in a new
CC   method for the in vitro activation of cytotoxic T cells (CTC).
CC   This is achieved by incubating the CTCs with antigen presenting
CC   cells loaded with an appropriate immunogenic peptide (e.g. one
CC   of the above peptides). By selecting the peptides used the
CC   following diseases and infections can be treated; cancer, AIDS,
CC   hepatitis, other viral and bacterial infections, malaria and
CC   tuberculosis.
SQ   Sequence   9 AA;

  Query Match           75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPTGHSY 35
           |||||||||
Qy      4 EADPTGHSY 12


RESULT   11
ID   R83932 standard; peptide; 9 AA.
AC   R83932;
DT   05-JUN-1996  (first entry)
DE   MHC class I restricted antigenic peptide #2.
KW   MHC class I; antigen; MAGE; melanoma; breast cancer; bladder cancer;
KW   Titermax; cytotoxic T-lymphocyte; tumour; pathogenic disease; bacteria;
KW   parasite; human; animal.
OS   Synthetic.
PN   WO9528958-A1.
PD   02-NOV-1995.
PF   21-APR-1995; U04975.
PR   22-APR-1994; US-233496.
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
```

PI Dyall R, Nikolic-Zugic J;
DR WPI; 95-382848/49.
PT Cytotoxic T-cell induction by MHC class I-restricted peptide in
PT adjuvant - useful for treating tumours and bacterial or parasitic
PT pathogenic diseases
PS Claim 11; Page 38; 50pp; English.
CC The sequences given in R83931-49 are MHC class I restricted 8-12
CC amino acid antigenic peptides. This peptide is derived from MAGE
CC and is present in melanoma, breast and bladder cancer. These
CC peptides may be administered to a subject in combination with a
CC suitable adjuvant, pref. Titermax (RTM), to induce cytotoxic T-
CC lymphocytes. This method may be used in the treatment of a tumour
CC or a pathogenic disease, esp. diseases of bacterial or parasitic
CC origin, in humans and animals, e.g monkeys, dogs cows, horses, etc.
SQ Sequence 9 AA;

Query Match 75.3%; Score 61; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.12e-17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 EADPTGHSY 35
||||||||||
Qy 4 EADPTGHSY 12

RESULT 12
ID R63675 standard; Protein; 9 AA.
AC R63675;
DT 22-JUN-1995 (first entry)
DE Synthetic peptide derived from exon 3.1 of MAGE 1.
KW Melanoma antigen-1; MAGE-1; cytolytic T cells; antigen E; exon 3.1.
OS Synthetic.
PN WO9423031-A.
PD 13-OCT-1994.
PF 17-MAR-1994; U02877.
PR 26-MAR-1993; US-037230.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-falleur T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;
DR WPI; 94-333192/41.
PT New tumour rejection antigen precursor MAGE3 - useful in
PT treatment and diagnosis of cancer
PS Example 34; Page 36; 105pp; English.
CC R63675 is a synthetic peptide derived from exon 3.1 of melanoma
CC antigen-1 (MAGE-1), it was used to transfer antigen-E cytolytic T
CC lymphocyte sensitivity to normally non-sensitive cells.
SQ Sequence 9 AA;

Query Match 75.3%; Score 61; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.12e-17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 EADPTGHSY 35
||||||||||
Qy 4 EADPTGHSY 12

RESULT 13
ID R75954 standard; Peptide; 9 AA.
AC R75954;
DT 06-MAR-1996 (first entry)
DE Melanoma antigen (MAGE-1) epitope.
KW MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;
KW cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.
OS Homo sapiens.
PN WO9519783-A1.
PD 27-JUL-1995.
PF 25-JAN-1995; U01000.
PR 25-JAN-1994; US-186266.
PA (CYTE-) CYTEL CORP.
PI Celis E, Grey HM, Kubo RT, Sette A;
DR WPI; 95-269270/35.
PT Immunogenic peptide(s) that induce immune response to cancer cells
PT - that express a MAGE-3 protein peptide epitope used in vaccines or
PT adoptive immuno:therapy to induce cytotoxic T lymphocytes
PS Example; Page 33; 44pp; English.
CC R75954 is derived from MAGE-1 protein. It was used to show the
CC specificity of CTL response to MAGE-3 peptides shown in R75942-53.
CC R75942 is derived from the sequence of the melanoma antigen (MAGE-3)
CC protein and can be used to elicit a primary cytotoxic T lymphocyte
CC response against cells expressing MAGE-3. Synthetic peptides R75945-53
CC can be used therapeutically to elicit CTL responses to melanoma, breast,
CC colon, prostate, or other cells which express proteins with this epitope.
CC The peptides have specific HLA-A1 binding capacity.
SQ Sequence 9 AA;

Query Match 75.3%; Score 61; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.12e-17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 EADPTGHSY 35
||||||||||
Qy 4 EADPTGHSY 12

RESULT 14
ID R29769 standard; Peptide; 9 AA.
AC R29769;
DT 22-APR-1993 (first entry)
DE Antigen E peptide.
KW Antigen; tumorigenic cell; A+ B+; T-cell; response; syngeneic;
KW animal; mouse; tumour rejection antigen precusor; TRAP; P1A.
OS Homo sapiens.
PN WO9220356-A.
PD 26-NOV-1992.
PF 22-MAY-1992; U04354.
PR 23-MAY-1991; US-705702.
PR 09-JUL-1991; US-728838.
PR 23-SEP-1991; US-764364.
PR 12-DEC-1991; US-807043.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;
PI Van Den Eynde B, Van Der Bruggen P, Van Pel A;
DR WPI; 92-415460/50.
PT Nucleic acid mol. encoding a human tumour rejection antigen
PT precursor - useful as an immunostimulant in a vaccine for
PT treating and preventing cancers, also useful in diagnosis
PS Disclosure; Page 97; 142pp; English.
CC This sequence represents the sequence of the antigen E. Antigens such
CC as this one cause a T-cell response to be elicited which transplanted
CC into a syngeneic animal, usually a mouse. This antigen is derived from
CC thecell line MEL3.1. See also Q32351.
SQ Sequence 9 AA;

Query Match 75.3%; Score 61; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.12e-17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 EADPTGHSY 35
||||||||||
Qy 4 EADPTGHSY 12

RESULT 15
ID R82988 standard; Peptide; 9 AA.
AC R82988;
DT 26-FEB-1996 (first entry)
DE P815 antigenic peptide.
KW P815 antigen; P1A antigen; cancer; vaccine.
OS Synthetic.
PN WO9523874-A1.
PD 08-SEP-1995.
PF 23-FEB-1995; U02203.
PR 01-MAR-1994; US-204727.
PR 10-MAR-1994; US-209172.
PR 01-SEP-1994; US-299849.
PR 30-NOV-1994; US-346774.

```
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  Brasseur F,  Chomez P,  De Plaen E;
PI   De Smet C,  Gaugler B, Lethe B,  Marchand M,  Patard J;
PI   Szikora J,  Van Den Eynde B,  Van Derbruggen P,  Weynants P;
DR   WPI; 95-320586/41.
PT   Determn. of cancerous condition(s) - using a nucleic acid as a
PT   primer to determine expression of a MAGE tumour rejection antigen
PT   precursor
PS   Example 13; Page 22; 121pp; English.
CC   Using the sequence of the P815A antigen precursor gene P1A
CC   (T01176), an antigenic peptide (R82988) which was A+B+ (i.e.
CC   characteristic of cells which express both A and B antigens) was
CC   produced.  The peptide lysed PO.HTR cells in the presence of
CC   cytolytic T lymphocyte cell lines, and may be useful as a vaccine
CC   component.
SQ   Sequence   9 AA;

  Query Match           75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPTGHSY 35
           |||||||||
Qy      4 EADPTGHSY 12


RESULT   16
ID   R90692 standard; peptide; 9 AA.
AC   R90692;
DT   31-JUL-1996  (first entry)
DE   Human leukocyte antigen (HLA-A1) presented peptide MZ2-E.
KW   Human leukocyte antigen; HLA-A1; MAGE-1 derived;
KW   blood mononuclear cell; BMC; CD8-beta+ cell; cytolytic T cell;
KW   CTL cell; treatment; tumour cell; diagnosis; assay;
KW   presented peptide.
OS   Synthetic.
PN   WO9535500-A1.
PD   28-DEC-1995.
PF   14-JUN-1995; U07559.
PR   17-JUN-1994; US-261541.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  Coulie P,  Van Der Bruggen P;
DR   WPI; 96-058510/06.
PT   Prodn. of specific cytolytic T cell sub-populations - by contacting
PT   blood mononuclear cells with specific peptide(s) and a population of
PT   CD8-beta(+) cells
PS   Claim 5; Page 19; 25pp; English.
CC   The present peptide is the human leukocyte antigen (HLA-A1), MAGE-1
CC   derived presented peptide, MZ2-E. By contacting a sample of blood
CC   mononuclear cells (BMC) with the peptide (which binds directly to
CC   HLA-A1 mols. on the surface of the BMC) and CD8-beta+ cells (which
CC   stimulate peptide/HLA-A1 complex specific CD8-beta+ cells), a
CC   peptide/HLA-A1 complex specific cytolytic T (CTL) cell
CC   subpopulation can be obtd. . The CTL cells obtd. can be
CC   administered to a patient to treat tumour cell related conditions,
CC   and can be used in diagnostic methods, e.g. in assays for the
CC   peptide/HLA-A1 complex.
SQ   Sequence   9 AA;

  Query Match           75.3%;  Score 61;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 1.12e-17;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPTGHSY 35
           |||||||||
Qy      4 EADPTGHSY 12


RESULT   17
ID   R99342 standard; peptide; 9 AA.
AC   R99342;
DT   22-APR-1997  (first entry)
DE   HLA binding nonapeptide #6.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Synthetic.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis  -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Claim 9; Page 30; 41pp; English.
CC   R99337-R99342 represent nonapeptides of the invention.  These sequences
CC   bind to a HLA molecule on a cell, and provoke lysis by cytolytic T cells
CC   (CTLs) specific for a complex of the HLA molecule and nonapeptide.  These
CC   sequences are based on regions of proteins in the MAGE family of human
CC   tumour rejection antigen precursors.  They are specifically based on
CC   fragments of the tumour rejection antigen (TRA) of MAGE-1.  The
CC   nonapeptides can be used diagnostically to identify tumours expressing a
CC   particular HLA molecule, or to identify cancer cells.  The peptides can
CC   also be used therapeutically, to induce a CTL response to tumours (where
CC   the peptides are optionally coupled to tumour-specific antibodies), or to
CC   induce a response by CTLs that are otherwise inactive.  These sequences
CC   may also be used to expand specific CTLs in vitro for later return to the
CC   patient, such as for treating melanoma.  Tumour cells can be identified
CC   by using DNA encoding these sequences as probes.  Non-human cells
CC   transformed with the HLA-A1 gene and a DNA sequence encoding one of these
CC   peptides, can be used to generate CTLs, or to detect the presence of CTLs
CC   in human samples.  The non-human transformed cells, when polytransformed,
CC   are universal effector cells, and can be used in vaccines, or for
CC   treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match           72.8%;  Score 59;  DB 1;  Length 35;
  Best Local Similarity 88.9%;  Pred. No. 3.99e-16;
  Matches    8;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPTGHAY 35
           |||||||:|
Qy      4 EADPTGHSY 12


RESULT   18
ID   R99339 standard; peptide; 9 AA.
AC   R99339;
DT   22-APR-1997  (first entry)
DE   HLA binding nonapeptide #3.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Synthetic.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis  -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Claim 9; Page 30; 41pp; English.
CC   R99337-R99342 represent nonapeptides of the invention.  These sequences
CC   bind to a HLA molecule on a cell, and provoke lysis by cytolytic T cells
CC   (CTLs) specific for a complex of the HLA molecule and nonapeptide.  These
CC   sequences are based on regions of proteins in the MAGE family of human
```

```
CC   tumour rejection antigen precursors.  They are specifically based on
CC   fragments of the tumour rejection antigen (TRA) of MAGE-1.  The
CC   nonapeptides can be used diagnostically to identify tumours expressing a
CC   particular HLA molecule, or to identify cancer cells.  The peptides can
CC   also be used therapeutically, to induce a CTL response to tumours (where
CC   the peptides are optionally coupled to tumour-specific antibodies), or to
CC   induce a response by CTLs that are otherwise inactive.  These sequences
CC   may also be used to expand specific CTLs in vitro for later return to the
CC   patient, such as for treating melanoma.  Tumour cells can be identified
CC   by using DNA encoding these sequences as probes.  Non-human cells
CC   transformed with the HLA-A1 gene and a DNA sequence encoding one of these
CC   peptides, can be used to generate CTLs, or to detect the presence of CTLs
CC   in human samples.  The non-human transformed cells, when polytransformed,
CC   are universal effector cells, and can be used in vaccines, or for
CC   treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match           71.6%;  Score 58;  DB 1;  Length 35;
  Best Local Similarity  88.9%;  Pred. No. 2.33e-15;
  Matches     8;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPAGHSY 35
          ||||:||||
Qy      4 EADPTGHSY 12


RESULT   19
ID   R99337 standard; peptide; 9 AA.
AC   R99337;
DT   22-APR-1997  (first entry)
DE   HLA binding nonapeptide #1.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Synthetic.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Claim 9; Page 30; 41pp; English.
CC   R99337-R99342 represent nonapeptides of the invention.  These sequences
CC   bind to a HLA molecule on a cell, and provoke lysis by cytolytic T cells
CC   (CTLs) specific for a complex of the HLA molecule and nonapeptide.  These
CC   sequences are based on regions of proteins in the MAGE family of human
CC   tumour rejection antigen precursors.  They are specifically based on
CC   fragments of the tumour rejection antigen (TRA) of MAGE-1.  The
CC   nonapeptides can be used diagnostically to identify tumours expressing a
CC   particular HLA molecule, or to identify cancer cells.  The peptides can
CC   also be used therapeutically, to induce a CTL response to tumours (where
CC   the peptides are optionally coupled to tumour-specific antibodies), or to
CC   induce a response by CTLs that are otherwise inactive.  These sequences
CC   may also be used to expand specific CTLs in vitro for later return to the
CC   patient, such as for treating melanoma.  Tumour cells can be identified
CC   by using DNA encoding these sequences as probes.  Non-human cells
CC   transformed with the HLA-A1 gene and a DNA sequence encoding one of these
CC   peptides, can be used to generate CTLs, or to detect the presence of CTLs
CC   in human samples.  The non-human transformed cells, when polytransformed,
CC   are universal effector cells, and can be used in vaccines, or for
CC   treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match           67.9%;  Score 55;  DB 1;  Length 35;
  Best Local Similarity 100.0%;  Pred. No. 4.16e-13;
  Matches     8;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     28 ADPTGHSY 35
          ||||||||
Qy      5 ADPTGHSY 12


RESULT   20
ID   R99340 standard; peptide; 9 AA.
AC   R99340;
DT   22-APR-1997  (first entry)
DE   HLA binding nonapeptide #4.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Synthetic.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Claim 9; Page 30; 41pp; English.
CC   R99337-R99342 represent nonapeptides of the invention.  These sequences
CC   bind to a HLA molecule on a cell, and provoke lysis by cytolytic T cells
CC   (CTLs) specific for a complex of the HLA molecule and nonapeptide.  These
CC   sequences are based on regions of proteins in the MAGE family of human
CC   tumour rejection antigen precursors.  They are specifically based on
CC   fragments of the tumour rejection antigen (TRA) of MAGE-1.  The
CC   nonapeptides can be used diagnostically to identify tumours expressing a
CC   particular HLA molecule, or to identify cancer cells.  The peptides can
CC   also be used therapeutically, to induce a CTL response to tumours (where
CC   the peptides are optionally coupled to tumour-specific antibodies), or to
CC   induce a response by CTLs that are otherwise inactive.  These sequences
CC   may also be used to expand specific CTLs in vitro for later return to the
CC   patient, such as for treating melanoma.  Tumour cells can be identified
CC   by using DNA encoding these sequences as probes.  Non-human cells
CC   transformed with the HLA-A1 gene and a DNA sequence encoding one of these
CC   peptides, can be used to generate CTLs, or to detect the presence of CTLs
CC   in human samples.  The non-human transformed cells, when polytransformed,
CC   are universal effector cells, and can be used in vaccines, or for
CC   treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match           67.9%;  Score 55;  DB 1;  Length 35;
  Best Local Similarity  88.9%;  Pred. No. 4.16e-13;
  Matches     8;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPTAHSY 35
          |||||:|||
Qy      4 EADPTGHSY 12


RESULT   21
ID   R99338 standard; peptide; 9 AA.
AC   R99338;
DT   22-APR-1997  (first entry)
DE   HLA binding nonapeptide #2.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Synthetic.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
```

```
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Claim 9; Page 30; 41pp; English.
CC   R99337-R99342 represent nonapeptides of the invention.  These sequences
CC   bind to a HLA molecule on a cell, and provoke lysis by cytolytic T cells
CC   (CTLs) specific for a complex of the HLA molecule and nonapeptide.  These
CC   sequences are based on regions of proteins in the MAGE family of human
CC   tumour rejection antigen precursors.  They are specifically based on
CC   fragments of the tumour rejection antigen (TRA) of MAGE-1.  The
CC   nonapeptides can be used diagnostically to identify tumours expressing a
CC   particular HLA molecule, or to identify cancer cells.  The peptides can
CC   also be used therapeutically, to induce a CTL response to tumours (where
CC   the peptides are optionally coupled to tumour-specific antibodies), or to
CC   induce a response by CTLs that are otherwise inactive.  These sequences
CC   may also be used to expand specific CTLs in vitro for later return to the
CC   patient, such as for treating melanoma.  Tumour cells can be identified
CC   by using DNA encoding these sequences as probes.  Non-human cells
CC   transformed with the HLA-A1 gene and a DNA sequence encoding one of these
CC   peptides, can be used to generate CTLs, or to detect the presence of CTLs
CC   in human samples.  The non-human transformed cells, when polytransformed,
CC   are universal effector cells, and can be used in vaccines, or for
CC   treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match           66.7%;  Score 54;  DB 1;  Length 35;
  Best Local Similarity  88.9%;  Pred. No. 2.26e-12;
  Matches     8;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADATGHSY 35
          |||:|||||
Qy      4 EADPTGHSY 12


RESULT   22
ID   W13255 standard; Peptide; 10 AA.
AC   W13255;
DT   29-OCT-1997  (first entry)
DE   MAGE-3 TRAP HLA-B44 motif.
KW   HLA-B44; tumour rejection antigen; TRAP; cancer; melanoma;
KW   MAGE-3; human; diagnosis; therapy.
OS   Homo sapiens.
PN   WO9710837-A1.
PD   27-MAR-1997.
PF   19-SEP-1996; U15078.
PR   21-SEP-1995; US-531864.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  Coulie P,  Herman J,  Van der Bruggen P;
DR   WPI; 97-202614/18.
PT   HLA-B44 molecule binding peptide(s) - useful to identify HLA-B44
PT   positive cells, and develop products for diagnosis and therapy of,
PT   e.g. cancer
PS   Claim 2; Page 36; 55pp; English.
CC   A peptide (W13255) corresponds to amino acids 167-176 of MAGE-3
CC   tumour rejection antigen precursor (TRAP).  It binds to HLA-B44
CC   positive cells and provokes lysis by cytotoxic T lymphocytes.  It
CC   can be used in competition assays to identify similarly active
CC   peptides.  HLA-B44 binding peptides (see also W13251-54 and W13256)
CC   can be used to identify HLA-B44 positive cells, and to develop
CC   products for the diagnosis and therapy (e.g. antibodies) of
CC   disorders in which the TRAP is expressed, e.g. cancer
CC   (particularly melanoma).
SQ   Sequence   10 AA;

  Query Match           63.0%;  Score 51;  DB 1;  Length 36;
  Best Local Similarity  63.6%;  Pred. No. 3.22e-10;
  Matches     7;  Conservative    1;  Mismatches   3;  Indels   0;  Gaps   0;

Db     26 VMEVDPIGHLY 36
          |:| || || |
Qy      2 VKEADPTGHSY 12


RESULT   23
ID   W16333 standard; Protein; 91 AA.
AC   W16333;
DT   05-SEP-1997  (first entry)
DE   Baboon MAGE-3 homologue D.
KW   MAGE-3; tumour antigen; vaccine; cancer; immunotherapy; therapy.
OS   Papio sp.
PN   WO9713858-A2.
PD   17-APR-1997.
PF   10-OCT-1996; U16319.
PR   12-OCT-1995; US-005117.
PA   (CHIR ) CHIRON CORP.
PI   Ralston RO,  Ring DB;
DR   WPI; 97-235894/21.
DR   N-PSDB; T63348.
PT   New isolated baboon MAGE-3 homologues - used for stimulating the
PT   immune system of humans to generate an immune response against
PT   MAGE-3 tumour antigen
PS   Example; Page 29; 36pp; English.
CC   Baboon MAGE-3 polypeptides A, C, D, B and F (W16331-35) are
CC   homologues of the human MAGE-3 tumour antigen (see also W16330),
CC   and are encoded by gene fragments (T63346-50) derived from baboon
CC   testis RNA.  The baboon polypeptides, expressed in host cells or
CC   in vivo, can be used to stimulate the immune system of humans.
CC   They can also be used to generate an immune response against cells
CC   expressing MAGE-3.  They are different enough to be recognised as
CC   foreign antigens more readily than human MAGE-3, but similar enough
CC   to enable the immune system to mount a response that will also
CC   recognise human MAGE-3.
SQ   Sequence   91 AA;

  Query Match           61.7%;  Score 50;  DB 1;  Length 117;
  Best Local Similarity  50.0%;  Pred. No. 1.61e-09;
  Matches     6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db     30 ELMEVDPVGHLY 41
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT   24
ID   W16332 standard; Protein; 88 AA.
AC   W16332;
DT   05-SEP-1997  (first entry)
DE   Baboon MAGE-3 homologue C.
KW   MAGE-3; tumour antigen; vaccine; cancer; immunotherapy; therapy.
OS   Papio sp.
FH   Key             Location/Qualifiers
FT   Misc_difference 18
FT   /label= Val, Asp
PN   WO9713858-A2.
PD   17-APR-1997.
PF   10-OCT-1996; U16319.
PR   12-OCT-1995; US-005117.
PA   (CHIR ) CHIRON CORP.
PI   Ralston RO,  Ring DB;
DR   WPI; 97-235894/21.
DR   N-PSDB; T63347.
PT   New isolated baboon MAGE-3 homologues - used for stimulating the
PT   immune system of humans to generate an immune response against
PT   MAGE-3 tumour antigen
PS   Example; Page 29; 36pp; English.
CC   Baboon MAGE-3 polypeptides A, C, D, B and F (W16331-35) are
CC   homologues of the human MAGE-3 tumour antigen (see also W16330),
CC   and are encoded by gene fragments (T63346-50) derived from baboon
CC   testis RNA.  The baboon polypeptides, expressed in host cells or
CC   in vivo, can be used to stimulate the immune system of humans.
CC   They can also be used to generate an immune response against cells
CC   expressing MAGE-3.  They are different enough to be recognised as
CC   foreign antigens more readily than human MAGE-3, but similar enough
CC   to enable the immune system to mount a response that will also
```

```
CC   recognise human MAGE-3.
SQ   Sequence   88 AA;

  Query Match           61.7%;  Score 50;  DB 1;  Length 122;
  Best Local Similarity  50.0%;  Pred. No. 1.61e-09;
  Matches     6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db     38 ELMEVDPVGHLY 49
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT   25
ID   W16335 standard; Protein; 91 AA.
AC   W16335;
DT   05-SEP-1997  (first entry)
DE   Baboon MAGE-3 homologue F.
KW   MAGE-3; tumour antigen; vaccine; cancer; immunotherapy; therapy.
OS   Papio sp.
FH   Key             Location/Qualifiers
FT   Misc_difference 77
FT   /label= Val, Phe
PN   WO9713858-A2.
PD   17-APR-1997.
PF   10-OCT-1996; U16319.
PR   12-OCT-1995; US-005117.
PA   (CHIR ) CHIRON CORP.
PI   Ralston RO,  Ring DB;
DR   WPI; 97-235894/21.
DR   N-PSDB; T63350.
PT   New isolated baboon MAGE-3 homologues - used for stimulating the
PT   immune system of humans to generate an immune response against
PT   MAGE-3 tumour antigen
PS   Example; Page 29; 36pp; English.
CC   Baboon MAGE-3 polypeptides A, C, D, B and F (W16331-35) are
CC   homologues of the human MAGE-3 tumour antigen (see also W16330),
CC   and are encoded by gene fragments (T63346-50) derived from baboon
CC   testis RNA.  The baboon polypeptides, expressed in host cells or
CC   in vivo, can be used to stimulate the immune system of humans.
CC   They can also be used to generate an immune response against cells
CC   expressing MAGE-3.  They are different enough to be recognised as
CC   foreign antigens more readily than human MAGE-3, but similar enough
CC   to enable the immune system to mount a response that will also
CC   recognise human MAGE-3.
SQ   Sequence   91 AA;

  Query Match           61.7%;  Score 50;  DB 1;  Length 125;
  Best Local Similarity  50.0%;  Pred. No. 1.61e-09;
  Matches     6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db     38 ELMEVDPVGHLY 49
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT   26
ID   W16330 standard; Protein; 126 AA.
AC   W16330;
DT   05-SEP-1997  (first entry)
DE   Human MAGE-3 tumour antigen.
KW   MAGE-3; tumour antigen; vaccine; cancer; immunotherapy; therapy.
OS   Homo sapiens.
PN   WO9713858-A2.
PD   17-APR-1997.
PF   10-OCT-1996; U16319.
PR   12-OCT-1995; US-005117.
PA   (CHIR ) CHIRON CORP.
PI   Ralston RO,  Ring DB;
DR   WPI; 97-235894/21.
DR   N-PSDB; T63345.
PT   New isolated baboon MAGE-3 homologues - used for stimulating the
PT   immune system of humans to generate an immune response against
PT   MAGE-3 tumour antigen
PS   Example; Page 29; 36pp; English.
CC   A polypeptide (W16330) comprises a partial sequence of human
CC   MAGE-3, a tumour antigen recognised on melanoma cells by
CC   autologous cytolytic T lymphocytes.  It shows homology to baboon
CC   polypeptides (W16311-35) encoded by gene fragments (T63346-50)
CC   that were derived from baboon testis RNA using primers (see also
CC   T63351-53) based on human MAGE-3 sequences.  The baboon MAGE-3
CC   homologues can be used to stimulate the immune system of humans
CC   to generate an immune response against MAGE-3.
SQ   Sequence   126 AA;

  Query Match           61.7%;  Score 50;  DB 1;  Length 152;
  Best Local Similarity  50.0%;  Pred. No. 1.61e-09;
  Matches     6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db     30 ELMEVDPIGHLY 41
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT   27
ID   W16331 standard; Protein; 125 AA.
AC   W16331;
DT   05-SEP-1997  (first entry)
DE   Baboon MAGE-3 homologue A.
KW   MAGE-3; tumour antigen; vaccine; cancer; immunotherapy; therapy.
OS   Papio sp.
FH   Key             Location/Qualifiers
FT   Misc_difference 18
FT   /label= Val, Asp
FT   Misc_difference 116
FT   /note= "residue 116 could not be determined due to
FT   DNA sequence ambiguity"
FT   Misc_difference 124
FT   /note= "residue 124 could not be determined due to
FT   DNA sequence ambiguity"
PN   WO9713858-A2.
PD   17-APR-1997.
PF   10-OCT-1996; U16319.
PR   12-OCT-1995; US-005117.
PA   (CHIR ) CHIRON CORP.
PI   Ralston RO,  Ring DB;
DR   WPI; 97-235894/21.
DR   N-PSDB; T63346.
PT   New isolated baboon MAGE-3 homologues - used for stimulating the
PT   immune system of humans to generate an immune response against
PT   MAGE-3 tumour antigen
PS   Example; Page 29; 36pp; English.
CC   Baboon MAGE-3 polypeptides A, C, D, B and F (W16331-35) are
CC   homologues of the human MAGE-3 tumour antigen (see also W16330),
CC   and are encoded by gene fragments (T63346-50) derived from baboon
CC   testis RNA.  The baboon polypeptides, expressed in host cells or
CC   in vivo, can be used to stimulate the immune system of humans.
CC   They can also be used to generate an immune response against cells
CC   expressing MAGE-3.  They are different enough to be recognised as
CC   foreign antigens more readily than human MAGE-3, but similar enough
CC   to enable the immune system to mount a response that will also
CC   recognise human MAGE-3.
SQ   Sequence   125 AA;

  Query Match           61.7%;  Score 50;  DB 1;  Length 159;
  Best Local Similarity  50.0%;  Pred. No. 1.61e-09;
  Matches     6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db     38 ELMEVDPIGHLY 49
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT   28
ID   W16334 standard; Protein; 126 AA.
AC   W16334;
DT   05-SEP-1997  (first entry)
```

```
DE   Baboon MAGE-3 homologue B.
KW   MAGE-3; tumour antigen; vaccine; cancer; immunotherapy; therapy.
OS   Papio sp.
FH   Key             Location/Qualifiers
FT   Misc_difference 77
FT   /label= Val, Phe
FT   Misc_difference 110
FT   /note= "residue 110 could not be determined due to
FT   DNA sequence ambiguity"
FT   Misc_difference 117
FT   /note= "residue 117 could not be determined due to
FT   DNA sequence ambiguity"
FT   Misc_difference 125
FT   /note= "residue 125 could not be determined due to
FT   DNA sequence ambiguity"
PN   WO9713858-A2.
PD   17-APR-1997.
PF   10-OCT-1996; U16319.
PR   12-OCT-1995; US-005117.
PA   (CHIR ) CHIRON CORP.
PI   Ralston RO,  Ring DB;
DR   WPI; 97-235894/21.
DR   N-PSDB; T63349.
PT   New isolated baboon MAGE-3 homologues - used for stimulating the
PT   immune system of humans to generate an immune response against
PT   MAGE-3 tumour antigen
PS   Example; Page 29; 36pp; English.
CC   Baboon MAGE-3 polypeptides A, C, D, B and F (W16331-35) are
CC   homologues of the human MAGE-3 tumour antigen (see also W16330),
CC   and are encoded by gene fragments (T63346-50) derived from baboon
CC   testis RNA.  The baboon polypeptides, expressed in host cells or
CC   in vivo, can be used to stimulate the immune system of humans.
CC   They can also be used to generate an immune response against cells
CC   expressing MAGE-3.  They are different enough to be recognised as
CC   foreign antigens more readily than human MAGE-3, but similar enough
CC   to enable the immune system to mount a response that will also
CC   recognise human MAGE-3.
SQ   Sequence   126 AA;

  Query Match           61.7%;  Score 50;  DB 1;  Length 160;
  Best Local Similarity  50.0%;  Pred. No. 1.61e-09;
  Matches    6;  Conservative    3;  Mismatches   3;  Indels   0;  Gaps   0;

Db     38 ELMEVDPVGHLY 49
          :::| || || |
Qy      1 DVKEADPTGHSY 12


RESULT   29
ID   R99341 standard; peptide; 9 AA.
AC   R99341;
DT   22-APR-1997  (first entry)
DE   HLA binding nonapeptide #5.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Synthetic.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Claim 9; Page 30; 41pp; English.
CC   R99337-R99342 represent nonapeptides of the invention.  These sequences
CC   bind to a HLA molecule on a cell, and provoke lysis by cytolytic T cells
CC   (CTLs) specific for a complex of the HLA molecule and nonapeptide.  These
CC   sequences are based on regions of proteins in the MAGE family of human
CC   tumour rejection antigen precursors.  They are specifically based on
CC   fragments of the tumour rejection antigen (TRA) of MAGE-1.  The
CC   nonapeptides can be used diagnostically to identify tumours expressing a
CC   particular HLA molecule, or to identify cancer cells.  The peptides can
CC   also be used therapeutically, to induce a CTL response to tumours (where
CC   the peptides are optionally coupled to tumour-specific antibodies), or to
CC   induce a response by CTLs that are otherwise inactive.  These sequences
CC   may also be used to expand specific CTLs in vitro for later return to the
CC   patient, such as for treating melanoma.  Tumour cells can be identified
CC   by using DNA encoding these sequences as probes.  Non-human cells
CC   transformed with the HLA-A1 gene and a DNA sequence encoding one of these
CC   peptides, can be used to generate CTLs, or to detect the presence of CTLs
CC   in human samples.  The non-human transformed cells, when polytransformed,
CC   are universal effector cells, and can be used in vaccines, or for
CC   treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match           60.5%;  Score 49;  DB 1;  Length 35;
  Best Local Similarity  88.9%;  Pred. No. 7.89e-09;
  Matches    8;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Db     27 EADPTGASY 35
          |||||| ||
Qy      4 EADPTGHSY 12


RESULT   30
ID   W13869 standard; Protein; 308 AA.
AC   W13869;
DT   12-MAY-1997  (first entry)
DE   Rad protein.
KW   Human; rad; C9D6; diabetogene; diabetes; obesity; ras/GTPase; antibody;
KW   glucose metabolism; Type II diabetes; transgenic animal.
OS   Homo sapiens.
PN   US5589374-A.
PD   31-DEC-1996.
PF   19-JUN-1992; 901710.
PR   19-JUN-1992; US-901710.
PR   11-JUN-1993; US-076089.
PA   (JOSL-) JOSLIN DIABETES CENT INC.
PI   Kahn CR,  Reynet C;
DR   WPI; 97-107140/10.
DR   N-PSDB; T60034.
PT   DNA encoding human rad protein - for prodn. of recombinant protein,
PT   useful to prepare antibodies and in therapeutic compsns.
PS   Claim 8; Column 17-20; 17pp; English.
CC   This sequence represents the human rad protein.  Rad (formerly referred
CC   to as C9D6) is a diabetogene.  A diabetogene is a gene whose expression,
CC   at the mRNA level, is altered in an individual with diabetes and/or
CC   obesity.  Rad encodes a member of the ras/GTPase related gene family, and
CC   shares 45-55% homology (at the nucleotide level) with other members of
CC   this family.  This sequence can be used for the production of recombinant
CC   rad, which can then be used to produce anti-rad antibodies, and in
CC   therapeutic compositions.  The rad protein can also be used for
CC   investigating the basis of glucose metabolism in normal and disease
CC   states at the level of gene structure and gene expression.  The protein
CC   can also be used to determine if a subject is at risk for a glucose
CC   metabolism related disorder (such as Type II diabetes), evaluating the
CC   effect of a treatment on the level of expression of a diabetogene,
CC   evaluating animal models for resemblance to a human disorder involving a
CC   diabetogene, and for making transgenic animals, and genetically altered
CC   cell lines for use in research.
SQ   Sequence   308 AA;

  Query Match           59.3%;  Score 48;  DB 1;  Length 334;
  Best Local Similarity  55.6%;  Pred. No. 3.78e-08;
  Matches    5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    145 EAEAAGHTY 153
          ||:::||:|
Qy      4 EADPTGHSY 12
```

```
RESULT   31
ID   R45431 standard; Protein; 308 AA.
AC   R45431;
DT   13-JUN-1994  (first entry)
DE   Diabetogene rad: A type II.
KW   Diabetogene rad; rad gene; diabetogene rad protein; diabetes;
KW   obesity.
OS   Homo sapiens.
PN   WO9400558-A.
PD   06-JAN-1994.
PF   11-JUN-1993; U05643.
PR   19-JUN-1992; US-901710.
PA   (JOSL-) JOSLIN DIABETES CENT INC.
PI   Kahn CR,  Reynet C;
DR   WPI; 94-026198/03.
DR   P-PSDB; Q55175.
PT   Purified DNA - includes type II diabetes-specific gene,
PT   Diabetogenic rad
PS   Disclosure; Page 37-38; 57pp; English.
CC   A gene encoding diabetogene rad protein R45431 is given in sequence
CC   Q55175.  The gene is used to determine the risk of diabetes or
CC   obesity, and the protein is used to treat deficiency of diabetogene
CC   product.
SQ   Sequence   308 AA;

  Query Match           59.3%;  Score 48;  DB 1;  Length 334;
  Best Local Similarity  55.6%;  Pred. No. 3.78e-08;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    145 EAEAAGHTY 153
          ||:::||:|
Qy      4 EADPTGHSY 12


RESULT   32
ID   R79148 standard; protein; 925 AA.
AC   R79148;
DT   28-DEC-1995  (first entry)
DE   Human insulin receptor tyrosine kinase inhibitor PC-1.
KW   Insulin receptro tyrosine kinase inhibitor; PC-1; insulin resistance;
KW   plasma cell membrane glycoprotein; insulin-dependent diabetes;
KW   diabetes mellitus.
OS   Homo sapiens.
PN   WO9519570-A.
PD   20-JUL-1995.
PF   28-DEC-1994; U14893.
PR   14-JAN-1994; US-182241.
PA   (GETH ) GENENTECH INC.
PA   (REGC ) UNIV CALIFORNIA.
PI   Goldfine ID,  Grupe A,  Maddux BA,  Spencer S,  Stewart TA;
DR   WPI; 95-263954/34.
PT   Diagnosis and treatment of insulin resistance - using antagonists,
PT   partic. antibodies to insulin receptor tyrosine kinase inhibitor
PT   proteins.
PS   Disclosure; Fig 1; 63pp; English.
CC   PC-1 is a class II (cytoplasmic N terminus) membrane glycoprotein.
CC   It is the same protein as liver nucleoside pyrophosphatase/alkaline
CC   phosphodiesterase I. Rebbe et al., Mol. Immuno., 30: 87-93 (1993).
CC   The size of PC-1 is 115-135 kDa, depending on the tissue studies. PC-
CC   1 also exists as a 230-260 kDa dimer. Human PC-1 has been deduced to
CC   have 873 AAs and is mapped to the chromosome location 6q22-6q23,
CC   Funakoshi et al., Arch. Biochem. Biophys., 295:180-187 (1992). The
CC   full length AA sequence of PC-1 is provided in R79148. PC-1 was
CC   purified from fibroblasts from an individual showing marked
CC   inhibition of insulin receptor autophosphorylation in vivo. A method
CC   for detecting or measuring the amount of PC-1 is claimed. This
CC   method can be used in the diagnosis and treatment of diseases
CC   involving inappropriate insulin receptor tyrosine kinase inhibitor
CC   expression e.g. non-insulin depedenent diabetes mellitus.
SQ   Sequence   925 AA;

  Query Match           59.3%;  Score 48;  DB 1;  Length 951;
  Best Local Similarity  66.7%;  Pred. No. 3.78e-08;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db    400 EPDSSGHSY 408
          |:|::||||
Qy      4 EADPTGHSY 12


RESULT   33
ID   P93191 standard; protein; 443 AA.
AC   P93191;
DT   26-MAR-1990  (first entry)
DE   Peptide with glutamine synthetase activity.
KW   Glutamine synthetase; Bacillus sp. EH113 (FERM BP-2281);
KW   E.coli DH1 pGS2 (FERM BP-2265); plasmid pGS2.
OS   Bacillus sp. EH113 (FERM BP-2281).
PN   DE3909249-A.
PD   02-NOV-1989.
PF   21-MAR-1989; 909249.
PR   23-MAR-1988; JP-P68671.
PA   (TOXN) Toyo Jozo Co. Ltd.
PI   Sagai H, Ohta H;
DR   WPI; 89-325536/45.
DR   N-PSDB; N92017.
PT   New DNA encoding new glutamine synthetase peptide - useful as reagent for
PT   ammonium ion assay, and transformed microorganisms providing efficient
PT   prodn. of high purity enzyme.
PS   Claim 2; page 11; 17pp; german.
CC   The peptide has glutamine synthetase activity.  It is useful as a reagent
CC   for assaying ammonium ions, or cpds. such as urea which generate them.
CC   DNA encoding the peptide is obtd. from Bacillus sp. EH113 (FERM BP-2281)
CC   and is expressed in host cell E.coli DH1 pGS2 (FERM BP-2265) using pGS2.
CC   This plasmid comprises a HindIII fragment from FERM BP-2281 inserted into
CC   HindIII-linearised pBR322.  Culturing the transformant gives efficient
CC   prodn. of high purity peptide at reduced costs.
SQ   Sequence   443 AA;

  Query Match           58.0%;  Score 47;  DB 1;  Length 471;
  Best Local Similarity  33.3%;  Pred. No. 1.77e-07;
  Matches     4;  Conservative    6;  Mismatches   2;  Indels   0;  Gaps   0;

Db    360 EVRSVDPAANPY 371
          :|:  ||::::|
Qy      1 DVKEADPTGHSY 12


RESULT   34
ID   R99350 standard; Protein; 9 AA.
AC   R99350;
DT   22-APR-1997  (first entry)
DE   MAGE-6 nonapeptide.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Homo sapiens.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
DR   N-PSDB; T35415.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis  -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Example 4; Fig 4; 41pp; English.
CC   R99343-R99350 represent MAGE nonapeptides, based on the tumour rejection
CC   antigen region of the full length MAGE sequences.  These peptides were
CC   used to design the nonapeptides of the invention (see R99337-R99342),
CC   which bind to a HLA molecule on a cell, and provoke lysis by cytolytic T
```

```
CC   cells (CTLs) specific for a complex of the HLA molecule and nonapeptide.
CC   The nonapeptides can be used diagnostically to identify tumours
CC   expressing a particular HLA molecule, or to identify cancer cells.  The
CC   peptides can also be used therapeutically, to induce a CTL response to
CC   tumours (where the peptides are optionally coupled to tumour-specific
CC   antibodies), or to induce a response by CTLs that are otherwise inactive.
CC   The peptide sequences may also be used to expand specific CTLs in vitro
CC   for later return to the patient, such as for treating melanoma.  Tumour
CC   cells can be identified by using DNA encoding the nonapeptides as probes.
CC   Non-human cells transformed with the HLA-A1 gene and a DNA sequence
CC   encoding one of the peptides, can be used to generate CTLs, or to detect
CC   the presence of CTLs in human samples.  The non-human transformed cells,
CC   when polytransformed, are universal effector cells, and can be used in
CC   vaccines, or for treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match           56.8%;  Score 46;  DB 1;  Length 35;
  Best Local Similarity  66.7%;  Pred. No. 8.04e-07;
  Matches    6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db     27 EVDPIGHVY 35
          | || || |
Qy      4 EADPTGHSY 12


RESULT   35
ID   R99349 standard; Protein; 9 AA.
AC   R99349;
DT   22-APR-1997  (first entry)
DE   MAGE-5/MAGE-51 nonapeptide.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Homo sapiens.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
DR   N-PSDB; T35414.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis  -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Example 4; Fig 4; 41pp; English.
CC   R99343-R99350 represent MAGE nonapeptides, based on the tumour rejection
CC   antigen region of the full length MAGE sequences.  These peptides were
CC   used to design the nonapeptides of the invention (see R99337-R99342),
CC   which bind to a HLA molecule on a cell, and provoke lysis by cytolytic T
CC   cells (CTLs) specific for a complex of the HLA molecule and nonapeptide.
CC   The nonapeptides can be used diagnostically to identify tumours
CC   expressing a particular HLA molecule, or to identify cancer cells.  The
CC   peptides can also be used therapeutically, to induce a CTL response to
CC   tumours (where the peptides are optionally coupled to tumour-specific
CC   antibodies), or to induce a response by CTLs that are otherwise inactive.
CC   The peptide sequences may also be used to expand specific CTLs in vitro
CC   for later return to the patient, such as for treating melanoma.  Tumour
CC   cells can be identified by using DNA encoding the nonapeptides as probes.
CC   Non-human cells transformed with the HLA-A1 gene and a DNA sequence
CC   encoding one of the peptides, can be used to generate CTLs, or to detect
CC   the presence of CTLs in human samples.  The non-human transformed cells,
CC   when polytransformed, are universal effector cells, and can be used in
CC   vaccines, or for treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match           56.8%;  Score 46;  DB 1;  Length 35;
  Best Local Similarity  66.7%;  Pred. No. 8.04e-07;
  Matches    6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db     27 EADPTSNTY 35
          |||||:::|
Qy      4 EADPTGHSY 12


Search completed: Tue Apr  7 08:43:12 1998
Job time : 8 secs.
```

****

****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 7 08:46:06 1998; MasPar time 1.50 Seconds
43.981 Million cell updates/sec
Tabular output not generated.

Title: >US-08-190-411A-4
Description: (1-12) from 5541104.pep
Perfect Score: 81
Sequence: 1 DVKEADPTGHSY 12

Scoring table: PAM 150
Gap 15

Searched: 60183 seqs, 5492030 residues

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database: a-issued
1:back1 2:51 3:52 4:53 5:54 6:55 7:56 8:57 9:PCT90
10:PCT91 11:PCT92 12:PCT93 13:PCT94 14:PCT95 15:PCT96

Statistics: Mean 15.516; Variance 38.247; scale 0.406

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	81	100.0	12	6	US-08-190-	Sequence 4, Applicatio	2.03e-04
2	61	75.3	9	14	PCT-US95-0	Sequence 1, Applicatio	1.60e-01
3	61	75.3	9	7	US-08-299-	Sequence 26, Applicati	1.60e-01
4	61	75.3	9	14	PCT-US95-0	Sequence 2, Applicatio	1.60e-01
5	61	75.3	9	7	US-08-186-	Sequence 1, Applicatio	1.60e-01
6	61	75.3	9	5	US-07-938-	Sequence 1, Applicatio	1.60e-01
7	61	75.3	9	7	US-08-443-	Sequence 12, Applicati	1.60e-01
8	61	75.3	9	5	US-08-073-	Sequence 12, Applicati	1.60e-01
9	55	67.9	8	5	US-08-073-	Sequence 14, Applicati	1.09e+00
10	55	67.9	8	7	US-08-443-	Sequence 14, Applicati	1.09e+00
11	49	60.5	8	5	US-08-073-	Sequence 13, Applicati	6.95e+00
12	49	60.5	8	7	US-08-443-	Sequence 13, Applicati	6.95e+00
13	49	60.5	8	5	US-07-938-	Sequence 21, Applicati	6.95e+00
14	49	60.5	417	6	US-08-430-	Sequence 6, Applicatio	6.95e+00
15	49	60.5	417	5	US-08-277-	Sequence 6, Applicatio	6.95e+00
16	49	60.5	417	3	US-07-649-	Sequence 6, Applicatio	6.95e+00
17	48	59.3	925	13	PCT-US94-1	Sequence 1, Applicatio	9.41e+00
18	46	56.8	9	5	US-08-073-	Sequence 20, Applicati	1.72e+01
19	46	56.8	9	5	US-07-938-	Sequence 9, Applicatio	1.72e+01
20	46	56.8	9	5	US-08-073-	Sequence 21, Applicati	1.72e+01
21	46	56.8	9	5	US-07-938-	Sequence 8, Applicatio	1.72e+01
22	46	56.8	9	7	US-08-443-	Sequence 21, Applicati	1.72e+01
23	46	56.8	9	7	US-08-443-	Sequence 22, Applicati	1.72e+01
24	46	56.8	9	5	US-07-938-	Sequence 7, Applicatio	1.72e+01
25	46	56.8	9	5	US-08-073-	Sequence 22, Applicati	1.72e+01
26	46	56.8	9	7	US-08-443-	Sequence 20, Applicati	1.72e+01
27	46	56.8	999	6	US-08-252-	Sequence 2, Applicatio	1.72e+01
28	43	53.1	9	14	PCT-US95-0	Sequence 1, Applicatio	4.15e+01
29	43	53.1	9	7	US-08-443-	Sequence 17, Applicati	4.15e+01
30	43	53.1	9	7	US-08-186-	Sequence 2, Applicatio	4.15e+01
31	43	53.1	9	5	US-07-938-	Sequence 4, Applicatio	4.15e+01
32	43	53.1	9	5	US-08-073-	Sequence 17, Applicati	4.15e+01
33	43	53.1	32	7	US-08-186-	Sequence 9, Applicatio	4.15e+01
34	43	53.1	559	12	PCT-US93-0	Sequence 12, Applicati	4.15e+01
35	43	53.1	559	8	US-08-368-	Sequence 12, Applicati	4.15e+01
36	43	53.1	2509	7	US-08-469-	Sequence 10, Applicati	4.15e+01
37	42	51.9	8	5	US-07-938-	Sequence 22, Applicati	5.54e+01
38	42	51.9	9	5	US-07-938-	Sequence 6, Applicatio	5.54e+01
39	42	51.9	9	5	US-08-073-	Sequence 19, Applicati	5.54e+01
40	42	51.9	9	7	US-08-443-	Sequence 19, Applicati	5.54e+01
41	42	51.9	375	12	PCT-US93-0	Sequence 8, Applicatio	5.54e+01
42	42	51.9	375	5	US-08-121-	Sequence 8, Applicatio	5.54e+01
43	41	50.6	36	7	US-08-050-	Sequence 30, Applicati	7.39e+01
44	41	50.6	91	6	US-08-225-	Sequence 16, Applicati	7.39e+01
45	41	50.6	343	7	US-08-187-	Sequence 2, Applicatio	7.39e+01
46	41	50.6	417	6	US-08-430-	Sequence 7, Applicatio	7.39e+01
47	41	50.6	417	5	US-08-277-	Sequence 7, Applicatio	7.39e+01
48	41	50.6	417	3	US-07-649-	Sequence 7, Applicatio	7.39e+01
49	41	50.6	459	1	5194375-6	Patent No. 5194375.	7.39e+01
50	41	50.6	530	7	US-08-187-	Sequence 4, Applicatio	7.39e+01
51	41	50.6	889	7	US-08-118-	Sequence 4, Applicatio	7.39e+01
52	41	50.6	942	5	US-08-141-	Sequence 14, Applicati	7.39e+01
53	41	50.6	942	8	US-08-541-	Sequence 14, Applicati	7.39e+01
54	40	49.4	414	7	US-08-332-	Sequence 46, Applicati	9.82e+01
55	40	49.4	414	7	US-08-332-	Sequence 52, Applicati	9.82e+01
56	40	49.4	414	6	US-08-188-	Sequence 52, Applicati	9.82e+01
57	40	49.4	430	5	US-08-147-	Sequence 25, Applicati	9.82e+01
58	40	49.4	430	11	PCT-US92-0	Sequence 19, Applicati	9.82e+01
59	40	49.4	430	12	PCT-US93-0	Sequence 9, Applicatio	9.82e+01
60	40	49.4	430	12	PCT-US93-0	Sequence 19, Applicati	9.82e+01
61	40	49.4	430	12	PCT-US93-1	Sequence 6, Applicatio	9.82e+01
62	40	49.4	430	12	PCT-US93-0	Sequence 4, Applicatio	9.82e+01
63	40	49.4	430	12	PCT-US93-0	Sequence 19, Applicati	9.82e+01
64	40	49.4	430	12	PCT-US93-0	Sequence 19, Applicati	9.82e+01
65	40	49.4	430	12	PCT-US93-0	Sequence 19, Applicati	9.82e+01
66	40	49.4	430	12	PCT-US93-0	Sequence 19, Applicati	9.82e+01
67	40	49.4	430	4	US-07-901-	Sequence 9, Applicatio	9.82e+01
68	40	49.4	430	8	US-08-447-	Sequence 25, Applicati	9.82e+01
69	40	49.4	430	7	US-08-155-	Sequence 19, Applicati	9.82e+01
70	40	49.4	430	7	US-08-479-	Sequence 6, Applicatio	9.82e+01
71	40	49.4	430	7	US-08-406-	Sequence 19, Applicati	9.82e+01
72	40	49.4	430	8	US-08-643-	Sequence 19, Applicati	9.82e+01
73	40	49.4	430	3	US-07-841-	Sequence 25, Applicati	9.82e+01
74	40	49.4	430	7	US-08-480-	Sequence 6, Applicatio	9.82e+01
75	40	49.4	430	9	PCT-US90-0	Sequence 2, Applicatio	9.82e+01
76	40	49.4	430	7	US-08-278-	Sequence 19, Applicati	9.82e+01
77	40	49.4	440	4	US-07-637-	Sequence 4, Applicatio	9.82e+01
78	40	49.4	440	4	US-08-112-	Sequence 4, Applicatio	9.82e+01
79	40	49.4	440	4	US-07-637-	Sequence 7, Applicatio	9.82e+01
80	40	49.4	824	7	US-08-221-	Sequence 3, Applicatio	9.82e+01
81	40	49.4	1833	14	PCT-US95-0	Sequence 18, Applicati	9.82e+01
82	39	48.1	9	5	US-07-938-	Sequence 5, Applicatio	1.30e+02
83	39	48.1	131	7	US-08-468-	Sequence 4, Applicatio	1.30e+02
84	39	48.1	181	7	US-07-949-	Sequence 4, Applicatio	1.30e+02
85	39	48.1	306	14	PCT-US95-0	Sequence 2, Applicatio	1.30e+02
86	39	48.1	350	15	PCT-US96-0	Sequence 2, Applicatio	1.30e+02
87	39	48.1	350	15	PCT-US96-0	Sequence 2, Applicatio	1.30e+02
88	39	48.1	440	4	US-08-112-	Sequence 5, Applicatio	1.30e+02
89	39	48.1	440	4	US-07-637-	Sequence 8, Applicatio	1.30e+02
90	39	48.1	528	12	PCT-US93-0	Sequence 13, Applicati	1.30e+02
91	39	48.1	628	14	PCT-US95-0	Sequence 6, Applicatio	1.30e+02
92	39	48.1	635	5	US-07-832-	Sequence 2, Applicatio	1.30e+02
93	39	48.1	878	14	PCT-US95-0	Sequence 2, Applicatio	1.30e+02
94	39	48.1	878	7	US-08-237-	Sequence 2, Applicatio	1.30e+02
95	39	48.1	1026	12	PCT-US93-1	Sequence 95, Applicati	1.30e+02

96	39	48.1	1026	14	PCT-US95-0	Sequence 95, Applicati	1.30e+02
97	39	48.1	1155	5	US-08-286-	Sequence 46, Applicati	1.30e+02
98	39	48.1	1203	14	PCT-US95-0	Sequence 103, Applicat	1.30e+02
99	39	48.1	1203	12	PCT-US93-1	Sequence 103, Applicat	1.30e+02
100	39	48.1	1513	12	PCT-US93-0	Sequence 2, Applicatio	1.30e+02

ALIGNMENTS

```
RESULT    1
ID   US-08-190-411A-4     STANDARD;     PRT;   12 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/08190411A.
XX
CC    Sequence 4, Application US/08190411A
CC    Patent No. 5541104
CC     GENERAL INFORMATION:
CC       APPLICANT:  Chen, Yao-Tseng; Stockert, Elisabeth;
CC       APPLICANT:  Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.;
CC       APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry;
CC       APPLICANT:  Old, Lloyd J.
CC       TITLE OF INVENTION:  MONOCLONAL ANTIBODIES WHICH BIND TO
CC       TITLE OF INVENTION:  TUMOR REJECTION ANTIGEN PRECURSOR MAGE-1, RECOMBI
  NANT MAGE-1,
CC       TITLE OF INVENTION:  AND MAGE-1 DERIVED IMMUNOGENIC PEPTIDES
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/190,411A
CC         FILING DATE:  01-FEBRUARY-1994
CC         CLASSIFICATION:  436
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  037,230
CC         FILING DATE:  26-MARCH-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/04354
CC         FILING DATE:  22-MAY-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/807,043
CC         FILING DATE:  12-DECEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/764,364
CC         FILING DATE:  23-SEPTEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/728,838
CC         APPLICATION NUMBER:  9-JULY-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/705,702
CC         FILING DATE:  23-MAY-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5541104man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5354
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  12 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   12 AA; 1318 MW; 944 CN;

  Query Match           100.0%;  Score 81;  DB 6;  Length 12;
  Best Local Similarity 100.0%;  Pred. No. 2.03e-04;
  Matches    12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 DVKEADPTGHSY 12
          ||||||||||||
Qy      1 DVKEADPTGHSY 12


RESULT    2
ID   PCT-US95-02121-1     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application PC/TUS9502121.
XX
CC    Sequence 1, Application PC/TUS9502121
CC     GENERAL INFORMATION:
CC       APPLICANT:
CC       TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR ELICITING
CC       TITLE OF INVENTION:  CTL IMMUNITY
CC       NUMBER OF SEQUENCES:  153
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US95/02121
CC         FILING DATE:  16-FEB-1995
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/197,484
CC         FILING DATE:  16-FEB-1994
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/935,811
CC         FILING DATE:  26-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/874,491
CC         FILING DATE:  27-APR-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/827,682
CC         FILING DATE:  29-JAN-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/749,568
CC         FILING DATE:  26-AUG-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Parmelee, Steven W.
CC         REGISTRATION NUMBER:  31,990
CC         REFERENCE/DOCKET NUMBER:  14137-26-4PC
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (206) 467-9600
CC         TELEFAX:  (415) 543-5043
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  unknown
CC         TOPOLOGY:  unknown
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match           75.3%;  Score 61;  DB 14;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.60e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Db      1 EADPTGHSY 9
          |||||||||
Qy      4 EADPTGHSY 12


RESULT    3
ID   US-08-299-849B-26     STANDARD;     PRT;    9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 26, Application US/08299849B.
XX
CC    Sequence 26, Application US/08299849B
CC    Patent No. 5612201
CC     GENERAL INFORMATION:
CC       APPLICANT:  De Plaen, Etienne; Boon-Falleur, Thierry;
CC       APPLICANT:  Leth , Bernard; Szikora, Jean-Pierre; De Smet, Charles;
CC       APPLICANT:  Chomez, Patrick
CC       TITLE OF INVENTION:  Isolated Nucleic Acid Molecules Useful In
CC       TITLE OF INVENTION:  Determining Expression Of A Tumor Antigen Precurs
  or
CC       NUMBER OF SEQUENCES:  48
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/299,849B
CC         FILING DATE:  1-SEPTEMBER-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/04354
CC         FILING DATE:  22-MAY-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/807,043
CC         FILING DATE:  12-DECEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/764,364
CC         FILING DATE:  23-SEPTEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/728,838
CC         APPLICATION NUMBER:  9-JULY-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/705,702
CC         FILING DATE:  23-May-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5612201man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5355
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  26:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acids
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match           75.3%;  Score 61;  DB 7;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.60e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      4 EADPTGHSY 12


RESULT    4
ID   PCT-US95-04975-2     STANDARD;     PRT;    9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application PC/TUS9504975.
XX
CC    Sequence 2, Application PC/TUS9504975
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nikolic-Zugic, Janko
CC       APPLICANT:  Dyall, Rubendra
CC       TITLE OF INVENTION:  INDUCTION OF CYTOTOXIC T LYMPHOCYTES (CTL)USING
CC       TITLE OF INVENTION:  ANTIGENIC PEPTIDES AND A SUITABLE ADJUVANT
CC       NUMBER OF SEQUENCES:  19
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cooper & Dunham LLP
CC         STREET:  1185 Avenue of the Americas
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10036
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US95/04975
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/233,496
CC         FILING DATE:  April 22, 1994
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  White Esq., John P.
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  45059/JPW/MSC/AMB
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  212-278-0400
CC         TELEFAX:  212-391-0525
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acids
CC         TOPOLOGY:  Linear
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  N
CC       ANTI-SENSE:  N
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match           75.3%;  Score 61;  DB 14;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.60e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      4 EADPTGHSY 12


RESULT    5
ID   US-08-186-266-1     STANDARD;     PRT;    9 AA.
XX
AC   xxxxxx
```

```
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application US/08186266.
XX
CC    Sequence 1, Application US/08186266
CC    Patent No. 5662907
CC     GENERAL INFORMATION:
CC       APPLICANT:  KUBO, Ralph T.
CC       APPLICANT:  GREY, Howard M.
CC       APPLICANT:  SETTE, Alessandro
CC       APPLICANT:  CELIS, Esteban
CC       TITLE OF INVENTION:  INDUCTION OF ANTI-TUMOR CYTOTOXIC
CC       TITLE OF INVENTION:  T LYMPHOCYTES IN HUMANS USING
CC       TITLE OF INVENTION:  SYNTHETIC PEPTIDE EPITOPES
CC       NUMBER OF SEQUENCES:  20
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Townsend and Townsend Khourie and Crew
CC         STREET:  Steuart Street Tower, One Market Plaza
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  US
CC         ZIP:  94105-1493
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/186,266
CC         FILING DATE:  25-JAN-1994
CC         CLASSIFICATION:  424
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/159,339
CC         FILING DATE:  29-NOV-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/103,396
CC         FILING DATE:  06-AUG-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/027,746
CC         FILING DATE:  05-MAR-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/926,666
CC         FILING DATE:  07-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Bastian, Kevin L.
CC         REGISTRATION NUMBER:  34,774
CC         REFERENCE/DOCKET NUMBER:  14137-50-4
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 543-9600
CC         TELEFAX:  (415) 543-5043
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match           75.3%;  Score 61;  DB 7;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.60e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      4 EADPTGHSY 12


RESULT    6
ID   US-07-938-334C-1     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application US/07938334C.
XX
CC    Sequence 1, Application US/07938334C
CC    Patent No. 5405940
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  De Plaen, Etienne; Lurquin Christophe; Traversari, Catia
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE GENES AND USES THEREOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/938,334C
CC         FILING DATE:  31-AUG-1992
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5405940man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 293
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  MAGE-1 derived nonapeptide
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match           75.3%;  Score 61;  DB 5;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.60e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      4 EADPTGHSY 12


RESULT    7
ID   US-08-443-341-12     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 12, Application US/08443341.
XX
CC    Sequence 12, Application US/08443341
CC    Patent No. 5695994
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
```

```
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/443,341
CC         FILING DATE:  17-MAY-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/073,103
CC         FILING DATE:  7-JUNE-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5695994man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.5
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match           75.3%;  Score 61;  DB 7;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.60e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHSY 9
           |||||||||
Qy       4 EADPTGHSY 12


RESULT    8
ID   US-08-073-103A-12     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 12, Application US/08073103A.
XX
CC    Sequence 12, Application US/08073103A
CC    Patent No. 5462871
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/073,103A
CC         FILING DATE:  7-JUNE-1993
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5462871man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match           75.3%;  Score 61;  DB 5;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.60e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHSY 9
           |||||||||
Qy       4 EADPTGHSY 12


RESULT    9
ID   US-08-073-103A-14     STANDARD;     PRT;   8 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 14, Application US/08073103A.
XX
CC    Sequence 14, Application US/08073103A
CC    Patent No. 5462871
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
```

```
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/073,103A
CC         FILING DATE:  7-JUNE-1993
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5462871man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  14:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  8 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   8 AA; 847 MW; 480 CN;

  Query Match           67.9%;  Score 55;  DB 5;  Length 8;
  Best Local Similarity 100.0%;  Pred. No. 1.09e+00;
  Matches     8;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 ADPTGHSY 8
          ||||||||
Qy      5 ADPTGHSY 12


RESULT   10
ID   US-08-443-341-14     STANDARD;     PRT;   8 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 14, Application US/08443341.
XX
CC    Sequence 14, Application US/08443341
CC    Patent No. 5695994
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/443,341
CC         FILING DATE:  17-MAY-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/073,103
CC         FILING DATE:  7-JUNE-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5695994man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.5
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  14:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  8 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   8 AA; 847 MW; 480 CN;

  Query Match           67.9%;  Score 55;  DB 7;  Length 8;
  Best Local Similarity 100.0%;  Pred. No. 1.09e+00;
  Matches     8;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 ADPTGHSY 8
          ||||||||
Qy      5 ADPTGHSY 12


RESULT   11
ID   US-08-073-103A-13     STANDARD;     PRT;   8 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 13, Application US/08073103A.
XX
CC    Sequence 13, Application US/08073103A
CC    Patent No. 5462871
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
```

```
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/073,103A
CC         FILING DATE:  7-JUNE-1993
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5462871man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  13:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  8 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   8 AA; 813 MW; 405 CN;

  Query Match           60.5%;  Score 49;  DB 5;  Length 8;
  Best Local Similarity 100.0%;  Pred. No. 6.95e+00;
  Matches     8;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHS 8
           ||||||||
Qy       4 EADPTGHS 11


RESULT   12
ID   US-08-443-341-13     STANDARD;     PRT;   8 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 13, Application US/08443341.
XX
CC    Sequence 13, Application US/08443341
CC    Patent No. 5695994
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/443,341
CC         FILING DATE:  17-MAY-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/073,103
CC         FILING DATE:  7-JUNE-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5695994man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.5
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  13:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  8 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   8 AA; 813 MW; 405 CN;

  Query Match           60.5%;  Score 49;  DB 7;  Length 8;
  Best Local Similarity 100.0%;  Pred. No. 6.95e+00;
  Matches     8;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHS 8
           ||||||||
Qy       4 EADPTGHS 11


RESULT   13
ID   US-07-938-334C-21     STANDARD;     PRT;   8 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 21, Application US/07938334C.
XX
CC    Sequence 21, Application US/07938334C
CC    Patent No. 5405940
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  De Plaen, Etienne; Lurquin Christophe; Traversari, Catia
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE GENES AND USES THEREOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/938,334C
CC         FILING DATE:  31-AUG-1992
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5405940man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 293
```

```
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  21:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  8 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   8 AA; 813 MW; 405 CN;

  Query Match           60.5%;  Score 49;  DB 5;  Length 8;
  Best Local Similarity 100.0%;  Pred. No. 6.95e+00;
  Matches     8;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHS 8
          ||||||||
Qy      4 EADPTGHS 11


RESULT   14
ID   US-08-430-787A-6     STANDARD;     PRT;   417 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application US/08430787A.
XX
CC    Sequence 6, Application US/08430787A
CC    Patent No. 5593674
CC     GENERAL INFORMATION:
CC       APPLICANT:  Drayna, Dennis T., Eaton, Dan L.
CC       TITLE OF INVENTION:  No. 5593674el Plasma Carboxypeptidase
CC       NUMBER OF SEQUENCES:  8
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Genentech, Inc.
CC         STREET:  460 Point San Bruno Blvd
CC         CITY:  South San Francisco
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  94080
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  patin (Genentech)
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/430,787A
CC         FILING DATE:  27-APR-1995
CC         CLASSIFICATION:  514
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/277,540
CC         FILING DATE:  19-JUL-1994
CC         APPLICATION NUMBER:  08/167727
CC         FILING DATE:  15-DEC-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/959944
CC         FILING DATE:  14-OCT-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/649591
CC         FILING DATE:  01-FEB-91
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hasak, Janet E.
CC         REGISTRATION NUMBER:  28,616
CC         REFERENCE/DOCKET NUMBER:  689D1C1D1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415/225-1896
CC         TELEFAX:  415/952-9881
CC         TELEX:  910/371-7168
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  417 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
SQ   SEQUENCE   417 AA; 48700 MW; 936226 CN;

  Query Match           60.5%;  Score 49;  DB 6;  Length 417;
  Best Local Similarity 66.7%;  Pred. No. 6.95e+00;
  Matches     8;  Conservative    0;  Mismatches   4;  Indels   0;  Gaps   0;

Db    105 DVKEDIPGRHSY 116
          ||||  |  |||
Qy      1 DVKEADPTGHSY 12


RESULT   15
ID   US-08-277-540-6     STANDARD;     PRT;   417 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application US/08277540.
XX
CC    Sequence 6, Application US/08277540
CC    Patent No. 5474901
CC     GENERAL INFORMATION:
CC       APPLICANT:  Drayna, Dennis T., Eaton, Dan L.
CC       TITLE OF INVENTION:  No. 5474901el Plasma Carboxypeptidase
CC       NUMBER OF SEQUENCES:  8
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Genentech, Inc.
CC         STREET:  460 Point San Bruno Blvd
CC         CITY:  South San Francisco
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  94080
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  patin (Genentech)
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/277,540
CC         FILING DATE:  19-JUL-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/167727
CC         FILING DATE:  15-DEC-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/959944
CC         FILING DATE:  14-OCT-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/649591
CC         FILING DATE:  01-FEB-91
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hasak, Janet E.
CC         REGISTRATION NUMBER:  28,616
CC         REFERENCE/DOCKET NUMBER:  689D1C1D1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415/225-1896
CC         TELEFAX:  415/952-9881
CC         TELEX:  910/371-7168
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  417 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
SQ   SEQUENCE   417 AA; 48700 MW; 936226 CN;

  Query Match           60.5%;  Score 49;  DB 5;  Length 417;
  Best Local Similarity 66.7%;  Pred. No. 6.95e+00;
  Matches     8;  Conservative    0;  Mismatches   4;  Indels   0;  Gaps   0;

Db    105 DVKEDIPGRHSY 116
```

```
          ||||  |  |||
Qy      1 DVKEADPTGHSY 12


RESULT   16
ID   US-07-649-591B-6     STANDARD;     PRT;   417 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 6, Application US/07649591B.
XX
CC    Sequence 6, Application US/07649591B
CC    Patent No. 5206161
CC     GENERAL INFORMATION:
CC       APPLICANT:  Dennis Drayna and Daniel Eaton
CC       TITLE OF INVENTION:  No. 5206161el Plasma Carboxypeptidase
CC       NUMBER OF SEQUENCES:  8
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Genentech, Inc.
CC         STREET:  460 Point San Bruno Blvd
CC         CITY:  South San Francisco
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  94080
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  patin (Genentech)
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/649,591B
CC         FILING DATE:  19910201
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hasak, Janet E.
CC         REGISTRATION NUMBER:  28,616
CC         REFERENCE/DOCKET NUMBER:  689
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  415/266-1896
CC         TELEFAX:  415/952-9881
CC         TELEX:  910/371-7168
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  417 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
SQ   SEQUENCE   417 AA; 48700 MW; 936226 CN;

  Query Match           60.5%;  Score 49;  DB 3;  Length 417;
  Best Local Similarity  66.7%;  Pred. No. 6.95e+00;
  Matches     8;  Conservative    0;  Mismatches   4;  Indels   0;  Gaps   0;

Db    105 DVKEDIPGRHSY 116
          ||||  |  |||
Qy      1 DVKEADPTGHSY 12


RESULT   17
ID   PCT-US94-14893-1     STANDARD;     PRT;   925 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application PC/TUS9414893.
XX
CC    Sequence 1, Application PC/TUS9414893
CC     GENERAL INFORMATION:
CC       APPLICANT:  Genentech, Inc.
CC       APPLICANT:  The Regents of the University of California
CC       APPLICANT:  Goldfine, Ira D.
CC       APPLICANT:  Grupe, Andrew
CC       APPLICANT:  Maddux, Betty A.
CC       APPLICANT:  Spencer, Steven
CC       APPLICANT:  Stewart, Timothy A.
CC       TITLE OF INVENTION:  Antagonists to Insulin Receptor Tyrosine Kinase I
  nhibitor
CC       NUMBER OF SEQUENCES:  1
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Genentech, Inc.
CC         STREET:  460 Point San Bruno Blvd
CC         CITY:  South San Francisco
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  94080
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  patin (Genentech)
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US94/14893
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/182241
CC         FILING DATE:  14-JAN-1994
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Kubinec, Jeffrey S.
CC         REGISTRATION NUMBER:  36,575
CC         REFERENCE/DOCKET NUMBER:  875P1PCT
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:
CC         TELEFAX:  415/952-9881
CC         TELEX:  910/371-7168
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  925 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
SQ   SEQUENCE   925 AA; 104924 MW; 4589090 CN;

  Query Match           59.3%;  Score 48;  DB 13;  Length 925;
  Best Local Similarity  66.7%;  Pred. No. 9.41e+00;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db    374 EPDSSGHSY 382
          |:|::||||
Qy      4 EADPTGHSY 12


RESULT   18
ID   US-08-073-103A-20     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 20, Application US/08073103A.
XX
CC    Sequence 20, Application US/08073103A
CC    Patent No. 5462871
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
```

```
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/073,103A
CC         FILING DATE:  7-JUNE-1993
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5462871man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  20:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 997 MW; 590 CN;

  Query Match           56.8%;  Score 46;  DB 5;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.72e+01;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTSNTY 9
           |||||:::|
Qy       4 EADPTGHSY 12


RESULT   19
ID   US-07-938-334C-9     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 9, Application US/07938334C.
XX
CC    Sequence 9, Application US/07938334C
CC    Patent No. 5405940
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  De Plaen, Etienne; Lurquin Christophe; Traversari, Catia
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE GENES AND USES THEREOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/938,334C
CC         FILING DATE:  31-AUG-1992
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5405940man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 293
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  9:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  MAGE-6 derived nonapeptide
SQ   SEQUENCE   9 AA; 1028 MW; 611 CN;

  Query Match           56.8%;  Score 46;  DB 5;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.72e+01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db       1 EVDPIGHVY 9
           | || || |
Qy       4 EADPTGHSY 12


RESULT   20
ID   US-08-073-103A-21     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 21, Application US/08073103A.
XX
CC    Sequence 21, Application US/08073103A
CC    Patent No. 5462871
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/073,103A
CC         FILING DATE:  7-JUNE-1993
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
```

```
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5462871man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  21:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 997 MW; 590 CN;

  Query Match           56.8%;  Score 46;  DB 5;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.72e+01;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTSNTY 9
           |||||:::|
Qy       4 EADPTGHSY 12


RESULT   21
ID   US-07-938-334C-8     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 8, Application US/07938334C.
XX
CC    Sequence 8, Application US/07938334C
CC    Patent No. 5405940
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  De Plaen, Etienne; Lurquin Christophe; Traversari, Catia
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE GENES AND USES THEREOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/938,334C
CC         FILING DATE:  31-AUG-1992
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5405940man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 293
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  8:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  MAGE-5 derived nonapeptide
SQ   SEQUENCE   9 AA; 997 MW; 590 CN;

  Query Match           56.8%;  Score 46;  DB 5;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.72e+01;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTSNTY 9
           |||||:::|
Qy       4 EADPTGHSY 12


RESULT   22
ID   US-08-443-341-21     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 21, Application US/08443341.
XX
CC    Sequence 21, Application US/08443341
CC    Patent No. 5695994
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/443,341
CC         FILING DATE:  17-MAY-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/073,103
CC         FILING DATE:  7-JUNE-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5695994man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.5
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  21:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
```

```
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 997 MW; 590 CN;

  Query Match           56.8%;  Score 46;  DB 7;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.72e+01;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTSNTY 9
          |||||:::|
Qy      4 EADPTGHSY 12


RESULT   23
ID   US-08-443-341-22     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 22, Application US/08443341.
XX
CC    Sequence 22, Application US/08443341
CC    Patent No. 5695994
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/443,341
CC         FILING DATE:  17-MAY-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/073,103
CC         FILING DATE:  7-JUNE-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5695994man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.5
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  22:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 1028 MW; 611 CN;

  Query Match           56.8%;  Score 46;  DB 7;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.72e+01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db      1 EVDPIGHVY 9
          | || || |
Qy      4 EADPTGHSY 12


RESULT   24
ID   US-07-938-334C-7     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 7, Application US/07938334C.
XX
CC    Sequence 7, Application US/07938334C
CC    Patent No. 5405940
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  De Plaen, Etienne; Lurquin Christophe; Traversari, Catia
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE GENES AND USES THEREOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/938,334C
CC         FILING DATE:  31-AUG-1992
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5405940man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 293
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  7:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  MAGE-5 derived nonapeptide
SQ   SEQUENCE   9 AA; 997 MW; 590 CN;

  Query Match           56.8%;  Score 46;  DB 5;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.72e+01;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTSNTY 9
          |||||:::|
Qy      4 EADPTGHSY 12


RESULT   25
```

```
ID   US-08-073-103A-22     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 22, Application US/08073103A.
XX
CC    Sequence 22, Application US/08073103A
CC    Patent No. 5462871
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/073,103A
CC         FILING DATE:  7-JUNE-1993
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5462871man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  22:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 1028 MW; 611 CN;

  Query Match           56.8%;  Score 46;  DB 5;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.72e+01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db      1 EVDPIGHVY 9
          | || || |
Qy      4 EADPTGHSY 12


RESULT   26
ID   US-08-443-341-20     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
```

```
DE   Sequence 20, Application US/08443341.
XX
CC    Sequence 20, Application US/08443341
CC    Patent No. 5695994
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/443,341
CC         FILING DATE:  17-MAY-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/073,103
CC         FILING DATE:  7-JUNE-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5695994man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.5
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  20:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 997 MW; 590 CN;

  Query Match           56.8%;  Score 46;  DB 7;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.72e+01;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTSNTY 9
          |||||:::|
Qy      4 EADPTGHSY 12


RESULT   27
ID   US-08-252-626A-2     STANDARD;     PRT;   999 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
```

```
DE   Sequence 2, Application US/08252626A.
XX
CC    Sequence 2, Application US/08252626A
CC    Patent No. 5585269
CC     GENERAL INFORMATION:
CC       APPLICANT:  Earp, Henry S.
CC       APPLICANT:  Graham, Douglas K.
CC       APPLICANT:  Dawson, Thomas L.
CC       APPLICANT:  Mullaney, David L.
CC       APPLICANT:  Snodgrass, Hiram R.
CC       TITLE OF INVENTION:  Isolated DNA Encoding C-MER
CC       TITLE OF INVENTION:  Protooncogene
CC       NUMBER OF SEQUENCES:  9
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Kenneth D. Sibley
CC         STREET:  P.O. Drawer 34009
CC         CITY:  Charlotte
CC         STATE:  No. 5585269th Carolina
CC         COUNTRY:  USA
CC         ZIP:  28234
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/252,626A
CC         FILING DATE:  02-JUN-1994
CC         CLASSIFICATION:  536
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Sibley, Kenneth D.
CC         REGISTRATION NUMBER:  31,665
CC         REFERENCE/DOCKET NUMBER:  5470-81
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (919) 881-3140
CC         TELEFAX:  (919) 881-3175
CC         TELEX:  575102
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC        LENGTH: 999 amino acids s
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   999 AA; 110390 MW; 5284465 CN;

  Query Match           56.8%;  Score 46;  DB 6;  Length 999;
  Best Local Similarity  63.6%;  Pred. No. 1.72e+01;
  Matches     7;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    320 QVKEADPLGNG 330
          :|||||| |::
Qy      1 DVKEADPTGHS 11


RESULT   28
ID   PCT-US95-04975-1     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 1, Application PC/TUS9504975.
XX
CC    Sequence 1, Application PC/TUS9504975
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nikolic-Zugic, Janko
CC       APPLICANT:  Dyall, Rubendra
CC       TITLE OF INVENTION:  INDUCTION OF CYTOTOXIC T LYMPHOCYTES (CTL)USING
CC       TITLE OF INVENTION:  ANTIGENIC PEPTIDES AND A SUITABLE ADJUVANT
CC       NUMBER OF SEQUENCES:  19
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cooper & Dunham LLP
CC         STREET:  1185 Avenue of the Americas
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10036
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US95/04975
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/233,496
CC         FILING DATE:  April 22, 1994
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  White Esq., John P.
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  45059/JPW/MSC/AMB
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  212-278-0400
CC         TELEFAX:  212-391-0525
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  Linear
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  N
CC       ANTI-SENSE:  N
SQ   SEQUENCE   9 AA; 1042 MW; 539 CN;

  Query Match           53.1%;  Score 43;  DB 14;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 4.15e+01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db      1 EVDPIGHLY 9
          | || || |
Qy      4 EADPTGHSY 12


RESULT   29
ID   US-08-443-341-17     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 17, Application US/08443341.
XX
CC    Sequence 17, Application US/08443341
CC    Patent No. 5695994
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
```

```
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/443,341
CC         FILING DATE:  17-MAY-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/073,103
CC         FILING DATE:  7-JUNE-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5695994man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.5
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  17:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 1042 MW; 539 CN;

  Query Match           53.1%;  Score 43;  DB 7;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 4.15e+01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db       1 EVDPIGHLY 9
           | || || |
Qy       4 EADPTGHSY 12


RESULT   30
ID   US-08-186-266-2     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 2, Application US/08186266.
XX
CC    Sequence 2, Application US/08186266
CC    Patent No. 5662907
CC     GENERAL INFORMATION:
CC       APPLICANT:  KUBO, Ralph T.
CC       APPLICANT:  GREY, Howard M.
CC       APPLICANT:  SETTE, Alessandro
CC       APPLICANT:  CELIS, Esteban
CC       TITLE OF INVENTION:  INDUCTION OF ANTI-TUMOR CYTOTOXIC
CC       TITLE OF INVENTION:  T LYMPHOCYTES IN HUMANS USING
CC       TITLE OF INVENTION:  SYNTHETIC PEPTIDE EPITOPES
CC       NUMBER OF SEQUENCES:  20
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Townsend and Townsend Khourie and Crew
CC         STREET:  Steuart Street Tower, One Market Plaza
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  US
CC         ZIP:  94105-1493
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/186,266
CC         FILING DATE:  25-JAN-1994
CC         CLASSIFICATION:  424
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/159,339
CC         FILING DATE:  29-NOV-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/103,396
CC         FILING DATE:  06-AUG-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/027,746
CC         FILING DATE:  05-MAR-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/926,666
CC         FILING DATE:  07-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Bastian, Kevin L.
CC         REGISTRATION NUMBER:  34,774
CC         REFERENCE/DOCKET NUMBER:  14137-50-4
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 543-9600
CC         TELEFAX:  (415) 543-5043
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   9 AA; 1042 MW; 539 CN;

  Query Match           53.1%;  Score 43;  DB 7;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 4.15e+01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db       1 EVDPIGHLY 9
           | || || |
Qy       4 EADPTGHSY 12


RESULT   31
ID   US-07-938-334C-4     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 4, Application US/07938334C.
XX
CC    Sequence 4, Application US/07938334C
CC    Patent No. 5405940
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  De Plaen, Etienne; Lurquin Christophe; Traversari, Catia
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE GENES AND USES THEREOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/938,334C
CC         FILING DATE:  31-AUG-1992
```

```
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5405940man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 293
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  MAGE-3 derived nonapeptide
SQ   SEQUENCE   9 AA; 1042 MW; 539 CN;

  Query Match           53.1%;  Score 43;  DB 5;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 4.15e+01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db      1 EVDPIGHLY 9
          | || || |
Qy      4 EADPTGHSY 12


RESULT   32
ID   US-08-073-103A-17     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 17, Application US/08073103A.
XX
CC    Sequence 17, Application US/08073103A
CC    Patent No. 5462871
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THER
  EOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/073,103A
CC         FILING DATE:  7-JUNE-1993
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5462871man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  17:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 1042 MW; 539 CN;

  Query Match           53.1%;  Score 43;  DB 5;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 4.15e+01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db      1 EVDPIGHLY 9
          | || || |
Qy      4 EADPTGHSY 12


RESULT   33
ID   US-08-186-266-9     STANDARD;     PRT;   32 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 9, Application US/08186266.
XX
CC    Sequence 9, Application US/08186266
CC    Patent No. 5662907
CC     GENERAL INFORMATION:
CC       APPLICANT:  KUBO, Ralph T.
CC       APPLICANT:  GREY, Howard M.
CC       APPLICANT:  SETTE, Alessandro
CC       APPLICANT:  CELIS, Esteban
CC       TITLE OF INVENTION:  INDUCTION OF ANTI-TUMOR CYTOTOXIC
CC       TITLE OF INVENTION:  T LYMPHOCYTES IN HUMANS USING
CC       TITLE OF INVENTION:  SYNTHETIC PEPTIDE EPITOPES
CC       NUMBER OF SEQUENCES:  20
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Townsend and Townsend Khourie and Crew
CC         STREET:  Steuart Street Tower, One Market Plaza
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  US
CC         ZIP:  94105-1493
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/186,266
CC         FILING DATE:  25-JAN-1994
CC         CLASSIFICATION:  424
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/159,339
CC         FILING DATE:  29-NOV-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/103,396
CC         FILING DATE:  06-AUG-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/027,746
CC         FILING DATE:  05-MAR-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/926,666
CC         FILING DATE:  07-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Bastian, Kevin L.
CC         REGISTRATION NUMBER:  34,774
CC         REFERENCE/DOCKET NUMBER:  14137-50-4
```

```
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 543-9600
CC         TELEFAX:  (415) 543-5043
CC     INFORMATION FOR SEQ ID NO:  9:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  32 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  1
CC         OTHER INFORMATION:  /note= "The Lys at the N-terminus
CC         OTHER INFORMATION:  is lipadated using two palmitic acid residues."
SQ   SEQUENCE   32 AA; 3365 MW; 5057 CN;

  Query Match           53.1%;  Score 43;  DB 7;  Length 32;
  Best Local Similarity  66.7%;  Pred. No. 4.15e+01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db     24 EVDPIGHLY 32
          | || || |
Qy      4 EADPTGHSY 12


RESULT   34
ID   PCT-US93-02172-12     STANDARD;     PRT;   559 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 12, Application PC/TUS9302172.
XX
CC    Sequence 12, Application PC/TUS9302172
CC     GENERAL INFORMATION:
CC       APPLICANT:  La Jolla Cancer Research Foundation
CC       TITLE OF INVENTION:  RECOMBINANT CALF INTESTINAL ALKALINE
CC       TITLE OF INVENTION:  PHOSPHATASE
CC       NUMBER OF SEQUENCES:  13
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  La Jolla Cancer Research Foundation
CC         STREET:  10901 North Torry Pines Road
CC         CITY:  La Jolla
CC         STATE:  California
CC         COUNTRY:  USA
CC         ZIP:  92037
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version 1.25 (EPO)
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US93/02172
CC         FILING DATE:  19930310
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/849,219
CC         FILING DATE:  10-MAR-1992
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619) 455-6480
CC         TELEFAX:  (619) 455-0181
CC         TELEX:
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  559 amino acids
CC         TYPE:  AMINO ACID
CC         TOPOLOGY:  linear
SQ   SEQUENCE   559 AA; 60255 MW; 1639274 CN;

  Query Match           53.1%;  Score 43;  DB 12;  Length 559;
  Best Local Similarity  41.7%;  Pred. No. 4.15e+01;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    427 NVTAAESSGSSY 438
          :|  |:::| ||
Qy      1 DVKEADPTGHSY 12


RESULT   35
ID   US-08-368-071-12     STANDARD;     PRT;   559 AA.
XX
AC   xxxxxx
XX
DT   01-JAN-1900
XX
DE   Sequence 12, Application US/08368071.
XX
CC    Sequence 12, Application US/08368071
CC    Patent No. 5707853
CC     GENERAL INFORMATION:
CC       APPLICANT:  MILLAN, JOSE L.
CC       TITLE OF INVENTION:  RECOMBINANT CALF INTESTINAL ALKALINE
CC       TITLE OF INVENTION:  PHOSPHATASE
CC       NUMBER OF SEQUENCES:  13
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  CAMPBELL AND FLORES
CC         STREET:  4370 LA JOLLA VILLAGE DRIVE, SUITE 700
CC         CITY:  SAN DIEGO
CC         STATE:  CALIFORNIA
CC         COUNTRY:  UNITED STATES
CC         ZIP:  92122
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/368,071
CC         FILING DATE:  03-JAN-1995
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  CAMPBELL, CATHRYN
CC         REGISTRATION NUMBER:  31,815
CC         REFERENCE/DOCKET NUMBER:  P-LJ 1275
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  619-535-9001
CC         TELEFAX:  619-535-8949
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  559 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
SQ   SEQUENCE   559 AA; 60255 MW; 1639274 CN;

  Query Match           53.1%;  Score 43;  DB 8;  Length 559;
  Best Local Similarity  41.7%;  Pred. No. 4.15e+01;
  Matches     5;  Conservative    4;  Mismatches   3;  Indels   0;  Gaps   0;

Db    427 NVTAAESSGSSY 438
          :|  |:::| ||
Qy      1 DVKEADPTGHSY 12


Search completed: Tue Apr  7 08:46:21 1998
Job time : 15 secs.
```